Maps of several variables of finite total variation and Helly-type selection principles[☆]

Vyacheslav V. Chistyakov[*,a], Yuliya V. Tretyachenko[a]

[a] *Department of Applied Mathematics and Informatics, State University Higher School of Economics, Bol'shaya Pechërskaya Street 25/12, Nizhny Novgorod 603155, Russia*

Abstract

Given two points $a=(a_1,\dots,a_n)$ and $b=(b_1,\dots,b_n)$ from $\mathbb{R}^n$ with $a<b$ componentwise and a map f from the rectangle $I_a^b=[a_1,b_1]\times\cdots\times[a_n,b_n]$ into a metric semigroup $M=(M,d,+)$, we study properties of the *total variation* $\mathrm{TV}(f,I_a^b)$ of f on I_a^b introduced by the first author in [V. V. Chistyakov, A selection principle for mappings of bounded variation of several variables, in: Real Analysis Exchange 27th Summer Symposium, 2003, 217–222], which extends the classical notion of C. Jordan's total variation ($n=1$) and the corresponding notions in the sense of [T. H. Hildebrandt, Introduction to the Theory of Integration, Academic Press, 1963] ($n=2$) and [A. S. Leonov, On the total variation for functions of several variables and a multidimensional analog of Helly's selection principle, Math. Notes 63 (1998) 61–71] ($n\in\mathbb{N}$) for real valued functions of n variables. The following Helly-type pointwise selection principle is proved: *If a sequence $\{f_j\}_{j\in\mathbb{N}}$ of maps from I_a^b into M is such that the closure in M of the set $\{f_j(x)\}_{j\in\mathbb{N}}$ is compact for each $x\in I_a^b$ and $C\equiv\sup_{j\in\mathbb{N}}\mathrm{TV}(f_j,I_a^b)$ is finite, then there exists a subsequence of $\{f_j\}_{j\in\mathbb{N}}$, which converges pointwise on I_a^b to a map f such that $\mathrm{TV}(f,I_a^b)\le C$.* A variant of this result is established concerning the weak pointwise convergence when values of maps lie in a reflexive Banach space $(M,\|\cdot\|)$ with separable dual M^*.

Key words: maps of several variables, total variation, selection principle,

[☆] The work of V. V. Chistyakov was supported by the State University Higher School of Economics Grant No. 09-08-0012 (Priority thematics "Investigation of Function Spaces").

[*] Corresponding author; tel.: +7 831 4169649; fax: +7 831 4169650.

Email addresses: `czeslaw@mail.ru` (Vyacheslav V. Chistyakov), `tretyachenko_y_v@mail.ru` (Yuliya V. Tretyachenko)

metric semigroup, pointwise convergence, weak convergence
MSC Classification: 26B30, 20M15, 28A20

1. Introduction

The classical Helly selection principle ([27]) states that *a bounded sequence of real valued functions on the closed interval, which is of uniformly bounded (Jordan) variation, contains a pointwise convergent subsequence whose limit is a function of bounded variation.* This theorem and its recent generalizations for real valued functions and metric space valued maps of one real variable ([7, 10, 15, 18, 19, 20, 21]) have numerous applications in different branches of Analysis (e.g., [6, 15, 25, 28, 33] and references therein).

Extensions of the Helly theorem for functions and maps of several real variables heavily depend upon notions of (bounded) variation used for these maps, which generalize different aspects of the classical Jordan variation of univariate functions and which are known to be quite numerous in the literature (e.g., [3, 8, 22, 24, 28, 30, 35, 38, 39, 41], and these references are far from being exhaustive on the topic). Under some approaches to the multidimensional variation ([2, 8, 34]) involving integration procedures Helly-type theorems are rather concerned with the *almost everywhere* convergence of extracted subsequences, and no stronger convergence can be expected in this case, but this convergence is far too weak for certain applications (such as those from [15]). On the other hand, there are definitions of the notion of variation for *real valued* functions of several variables ([28, 32]), which go back to Vitali [38], Hardy [26] and Krause [1, 22], such that a complete analogue of the Helly theorem holds with respect to the *pointwise* convergence of extracted subsequences. These counterparts of Helly's theorem are based on the notion of a (totally) *monotone* real valued function of several variables [9, 28, 40] and an appropriate generalization of Jordan's decomposition theorem when a function of bounded variation is represented as the difference of two monotone functions.

The aim of this paper is twofold. First, we study properties of the *total variation* of metric semigroup valued maps of several variables in the approach of Vitali, Hardy and Krause introduced by the first author in [14], which extends the classical notion of Jordan's total variation for maps of one variable and the notions of the total variation in the sense of Hildebrandt [28] for real valued functions of two variables and Leonov [32] for real valued functions of any finite number of variables. Second, we present two variants of

a Helly-type *pointwise* selection principle for metric semigroup valued maps and maps with values in a reflexive separable Banach space. The main difficulty that we overcome is that for metric semigroup valued maps there is no counterpart of Jordan's decomposition theorem, and we have to develop a completely different technique, whose two-dimesional variant is given in [5].

The paper is organized as follows. In Section 2 we present necessary definitions and our two main results, Theorems 1 and 2. In order to get to the proofs of these results as quick as possible, in Section 3 we collect all main ingredients and auxiliary known facts needed for their proofs. In Section 4 we prove Theorems 1 and 2. The remaining Sections 5–8 contain proofs of the results exposed in Section 3 and used in the proofs of the main theorems.

2. Definitions and main results

Throughout the paper we adopt and follow the Vitali-Hardy-Krause approach to the notion of variation for maps of several variables in the multiindex notation initiated in [12, 14] and developed in detail in [17] (equivalent approaches in different notation for real functions can be found in [31, 32]).

Let $\mathbb{N}$ and $\mathbb{N}_0$ stand for the sets of positive and nonnegative integers, respectively, and $n \in \mathbb{N}$. Given $x, y \in \mathbb{R}^n$, we write $x = (x_1, \ldots, x_n) = (x_i : i \in \{1, \ldots, n\})$ for the coordinate representation of x, and set $x + y = (x_1 + y_1, \ldots, x_n + y_n)$, and $x - y$ is defined similarly. The inequality $x < y$ will be understood componentwise, i.e., $x_i < y_i$ for all $i \in \{1, \ldots, n\}$, and a similar meaning applies to $x = y$, $x \le y$, $y \ge x$ and $y > x$. If $x < y$ or $x \le y$, we denote by I_x^y the rectangle $\prod_{i=1}^n [x_i, y_i] = [x_1, y_1] \times \cdots \times [x_n, y_n]$. Elements of the set $\mathbb{N}_0^n$ are as usual said to be *multiindices* and denoted by Greek letters and, given $\theta = (\theta_1, \ldots, \theta_n) \in \mathbb{N}_0^n$ and $x \in \mathbb{R}^n$, we set $|\theta| = \theta_1 + \cdots + \theta_n$ (the order of θ) and $\theta x = (\theta_1 x_1, \ldots, \theta_n x_n)$. The n-dimensional multiindices $0_n = (0, \ldots, 0)$ and $1_n = (1, \ldots, 1)$ will be denoted simply by 0 and 1, respectively (actually, the dimension of 0 and 1 will be clear from the context). We also put $\mathcal{E}(n) = \{\theta \in \mathbb{N}_0^n : \theta \le 1 \text{ and } |\theta| \text{ is even}\}$ (the set of 'even' multiindices) and $\mathcal{O}(n) = \{\theta \in \mathbb{N}_0^n : \theta \le 1 \text{ and } |\theta| \text{ is odd}\}$ (the set of 'odd' multiindices). For elements from the set $\mathcal{A}(n) = \{\alpha \in \mathbb{N}_0^n : 0 \ne \alpha \le 1\}$ we simply write $0 \ne \alpha \le 1$.

The domain of (almost) all maps under consideration will be a rectangle I_a^b with fixed $a, b \in \mathbb{R}^n$, $a < b$, called the *basic rectangle*. The range of maps will be a *metric semigroup* $(M, d, +)$, i.e., (M, d) is a metric space, $(M, +)$ is an Abelian semigroup with the operation of addition $+$, and d

is translation invariant: $d(u,v) = d(u+w, v+w)$ for all $u, v, w \in M$. A nontrivial example of a metric semigroup is as follows ([23, 36]): Let $(X, \|\cdot\|)$ be a real normed space and M be the family of all nonempty closed bounded convex subsets of X equipped with the Hausdorff metric d given by $d(U,V) = \max\{\mathrm{e}(U,V), \mathrm{e}(V,U)\}$, where $U, V \in M$ and $\mathrm{e}(U,V) = \sup_{u\in U} \inf_{v\in V} \|u-v\|$. Given $U, V \in M$, defining $U \oplus V$ as the closure in X of the Minkowski sum $\{u+v : u \in U, v \in V\}$ we find that the triple $(M, d, \oplus)$ is a metric semigroup.

Given $f : I_a^b \to (M, d, +)$, we define the *Vitali-type n-th mixed 'difference'* of f on a subrectangle $I_x^y \subset I_a^b$, where $x, y \in I_a^b$ and $x < y$, by (cf. [14])

$$\mathrm{md}_n(f, I_x^y) = d\Big(\sum_{\theta\in\mathcal{E}(n)} f\big(x + \theta\,(y-x)\big), \sum_{\eta\in\mathcal{O}(n)} f\big(x + \eta\,(y-x)\big)\Big). \tag{2.1}$$

For example, for the first three dimensions we have: if $n = 1$, then $\mathcal{E}(1) = \{0\}$ and $\mathcal{O}(1) = \{1\}$, and so, $\mathrm{md}_1(f, I_x^y) = d(f(x), f(y))$; if $n = 2$, then $\mathcal{E}(2) = \{(0,0), (1,1)\}$ and $\mathcal{O}(2) = \{(0,1), (1,0)\}$, and so,

$$\mathrm{md}_2(f, I_x^y) = d\big(f(x_1,x_2) + f(y_1,y_2), f(x_1,y_2) + f(y_1,x_2)\big);$$

if $n = 3$, then $\mathcal{E}(3) = \{(0,0,0), (1,1,0), (1,0,1), (0,1,1)\}$ and $\mathcal{O}(3) = \{(1,1,1), (1,0,0), (0,1,0), (0,0,1)\}$, and so,

$$\begin{aligned}\mathrm{md}_3(f, I_x^y) \;=\; d\Big(&f(x_1,x_2,x_3) + f(y_1,y_2,x_3) + f(y_1,x_2,y_3) + f(x_1,y_2,y_3),\\ &f(y_1,y_2,y_3) + f(y_1,x_2,x_3) + f(x_1,y_2,x_3) + f(x_1,x_2,y_3)\Big)\end{aligned}$$

(one may draw corresponding pictures to see the points where f is evaluated at the left and right hand places of d('left', 'right')).

Remark 2.1. Formally, the value $\mathrm{md}_n(f, I_x^y)$ from (2.1) is defined for $x < y$. Now if $x, y \in I_a^b$, $x \le y$ and $x \not< y$, then the right-hand side in (2.1) is equal to zero for any map $f : I_a^b \to M$. In fact, if $x_i = y_i$ for some $i \in \{1, \ldots, n\}$, then

$$\sum_{\theta\in\mathcal{E}(n)} f(x + \theta(y-x)) = \sum_{\overline{\theta}\in\mathcal{O}(n)} f(x + \overline{\theta}(y-x)).$$

In order to see this, given $\theta = (\theta_1, \ldots, \theta_n) \in \mathcal{E}(n)$, we set $\overline{\theta} = (\overline{\theta}_1, \ldots, \overline{\theta}_n) = (\theta_1, \ldots, \theta_{i-1}, 1-\theta_i, \theta_{i+1}, \ldots, \theta_n)$ and note that $\overline{\theta} \in \mathcal{O}(n)$ and, moreover, the map $\theta \mapsto \overline{\theta}$ is a bijection between $\mathcal{E}(n)$ and $\mathcal{O}(n)$. It remains to take into account that $x + \theta(y-x) = x + \overline{\theta}(y-x)$ for all $\theta \in \mathcal{E}(n)$, because

$$x_i + \theta_i(y_i - x_i) = x_i = x_i + (1-\theta_i)(y_i - x_i) = x_i + \overline{\theta}_i(y_i - x_i).$$

The *Vitali-type n-th variation* ([17, 32, 38]) of $f : I_a^b \to M$ is defined by

$$V_n(f, I_a^b) = \sup_{\mathcal{P}} \sum_{1\le\sigma\le\kappa} \mathrm{md}_n\big(f, I_{x[\sigma-1]}^{x[\sigma]}\big), \tag{2.2}$$

the supremum being taken over all multiindices $\kappa \in \mathbb{N}^n$ and all *net partitions* of I_a^b of the form $\mathcal{P} = \{x[\sigma]\}_{\sigma=0}^{\kappa}$, where points $x[\sigma] = (x_1(\sigma_1), \dots, x_n(\sigma_n))$ from I_a^b are indexed by $\sigma = (\sigma_1, \dots, \sigma_n) \in \mathbb{N}_0^n$ with $\sigma \le \kappa$ and satisfy the conditions: $x[0] = a$, $x[\kappa] = b$ and $x[\sigma - 1] < x[\sigma]$ for all $1 \le \sigma \le \kappa$ (in other words, a net partition $\mathcal{P}$ is the Cartesian product of ordinary partitions of closed intervals $[a_i, b_i]$, $i = 1, \dots, n$). Note that all rectangles $I_{x[\sigma-1]}^{x[\sigma]}$ of a net partition are non-degenerated, non-overlapping and their union is I_a^b.

In order to define the notion of the total variation of a map $f : I_a^b \to M$ we need the notion of variation of f of order less than n. Following [17], we define the *truncation of a point* $x \in \mathbb{R}^n$ *by a multiindex* $0 \ne \alpha \le 1$ by $x\lfloor\alpha = (x_i \,:\, i \in \{1, \dots, n\},\ \alpha_i = 1)$, and set $I_a^b\lfloor\alpha = I_{a\lfloor\alpha}^{b\lfloor\alpha}$. Clearly, $x\lfloor 1 = x$ and $I_a^b\lfloor 1 = I_a^b$, and if $x \in I_a^b$, then $x\lfloor\alpha \in I_a^b\lfloor\alpha$. For example, if $x = (x_1, x_2, x_3, x_4)$ and $\alpha = (1, 0, 0, 1)$, we have $x\lfloor\alpha = (x_1, x_4)$ and $I_a^b\lfloor\alpha = [a_1, b_1] \times [a_4, b_4]$. Given $f : I_a^b \to M$ and $z \in I_a^b$, we define the *truncated map* $f_\alpha^z : I_a^b\lfloor\alpha \to M$ *with the base at* z by $f_\alpha^z(x\lfloor\alpha) = f\big(z + \alpha\,(x - z)\big)$ for all $x \in I_a^b$. It follows that f_α^z depends only on $|\alpha|$ variables $x_i \in [a_i, b_i]$, for which $\alpha_i = 1$, and the other variables remain fixed and equal to z_j when $\alpha_j = 0$. In the above example we get $f_\alpha^z(x_1, x_4) = f_\alpha^z(x\lfloor\alpha) = f(x_1, z_2, z_3, x_4)$ for $(x_1, x_4) \in [a_1, b_1] \times [a_4, b_4]$.

Now, given $f : I_a^b \to M$ and $0 \ne \alpha \le 1$, the function $f_\alpha^a : I_a^b\lfloor\alpha \to M$ with the base at $z = a$ depends only on $|\alpha|$ variables, and so, making use of the definitions (2.2) and (2.1) with n replaced by $|\alpha|$, f replaced by f_α^a and I_a^b replaced by $I_a^b\lfloor\alpha$, we get the notion of the (*Hardy-Krause-type* [1, 22, 26]) *$|\alpha|$-th variation* of f, which is denoted by $V_{|\alpha|}(f_\alpha^a, I_a^b\lfloor\alpha)$.

The *total variation* of $f : I_a^b \to M$ in the sense of Hildebrandt ([13, 16], [28, III.6.3], [29] if $n = 2$) and Leonov ([12, 14, 17, 32] if $n \in \mathbb{N}$) is defined by

$$\mathrm{TV}(f, I_a^b) = \sum_{0\ne\alpha\le1} V_{|\alpha|}(f_\alpha^a, I_a^b\lfloor\alpha) \equiv \sum_{i=1}^{n} \sum_{\alpha\le1,\,|\alpha|=i} V_i(f_\alpha^a, I_a^b\lfloor\alpha), \tag{2.3}$$

the summations here and throughout the paper being taken over *n-dimensional* multiindices in the ranges specified under the summation sign.

For the first three dimensions $n = 1, 2, 3$ we have, respectively,

$$
\begin{aligned}
\mathrm{TV}(f, I_a^b) &= V_a^b(f), \quad \text{the usual Jordan variation on the interval } [a,b],\\
\mathrm{TV}(f, I_a^b) &= V_{a_1}^{b_1}\bigl(f(\,\cdot\,, a_2)\bigr) + V_{a_2}^{b_2}\bigl(f(a_1, \,\cdot\,)\bigr) + V_2(f, I_{a_1,a_2}^{b_1,b_2}),\\
\mathrm{TV}(f, I_a^b) &= V_{a_1}^{b_1}\bigl(f(\,\cdot\,, a_2, a_3)\bigr) + V_{a_2}^{b_2}\bigl(f(a_1, \,\cdot\,, a_3)\bigr) + V_{a_3}^{b_3}\bigl(f(a_1, a_2, \,\cdot\,)\bigr)\\
&\quad + V_2\bigl(f(\,\cdot\,, \,\cdot\,, a_3), I_{a_1,a_2}^{b_1,b_2}\bigr) + V_2\bigl(f(\,\cdot\,, a_2, \,\cdot\,), I_{a_1,a_3}^{b_1,b_3}\bigr)\\
&\quad + V_2\bigl(f(a_1, \,\cdot\,, \,\cdot\,), I_{a_2,a_3}^{b_2,b_3}\bigr) + V_3(f, I_{a_1,a_2,a_3}^{b_1,b_2,b_3}).
\end{aligned}
$$

We denote by $\mathrm{BV}(I_a^b; M)$ the space of all maps $f : I_a^b \to M$ of *finite* (or bounded) *total variation* (2.3).

Recall that a sequence $\{f_j\} \equiv \{f_j\}_{j\in\mathbb{N}}$ of maps from I_a^b into M is said: (a) to *converge pointwise* on I_a^b to a map $f : I_a^b \to M$ if $d(f_j(x), f(x)) \to 0$ as $j \to \infty$ for all $x \in I_a^b$; (b) to be *pointwise precompact* (on I_a^b) provided the closure in M of the set $\{f_j(x)\}_{j\in\mathbb{N}}$ is compact for all $x \in I_a^b$.

Our first main result, to be proved in Section 4, is the following Helly-type pointwise selection principle in the space $\mathrm{BV}(I_a^b; M)$:

Theorem 1. *A pointwise precompact sequence $\{f_j\}$ of maps from the rectangle I_a^b into a metric semigroup $(M, d, +)$ such that*

$$C \equiv \sup_{j\in\mathbb{N}} \mathrm{TV}(f_j, I_a^b) \quad \textit{is finite} \tag{2.4}$$

contains a subsequence which converges pointwise on I_a^b to a map $f \in \mathrm{BV}(I_a^b; M)$ such that $\mathrm{TV}(f, I_a^b) \le C$.

This result was announced in [14]. It contains as particular cases the results of [28, III.6.5] and [29] ($n = 2$ and $M = \mathbb{R}$), [32] ($n \in \mathbb{N}$ and $M = \mathbb{R}$) and [5] ($n = 2$ and M is a metric semigroup).

Our second main result (Theorem 2 below) is concerned with a weak analogue of Theorem 1 taking into account certain specific features when the values of maps under consideration lie in a reflexive separable Banach space.

Let $(M, \|\cdot\|)$ be a normed linear space over the field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ and M^* be its *dual*, i.e., $M^* = \mathrm{L}(M; \mathbb{K})$, the space of all continuous linear functionals on M. It is well-known that M^* is a Banach space under the norm $\|u^*\| = \sup\{|u^*(u)| : u \in M \text{ and } \|u\| \le 1\}$, $u^* \in M^*$. The natural duality between M and M^* is determined by the bilinear functional $\langle\cdot,\cdot\rangle : M \times M^* \to \mathbb{K}$ defined by $\langle u, u^*\rangle = u^*(u)$ for all $u \in M$ and $u^* \in M^*$, so that $|\langle u, u^*\rangle| \le$

$\|u\| \cdot \|u^*\|$, where $|\cdot|$ is the absolute value in $\mathbb{K}$. Recall that a sequence $\{u_j\} \subset M$ *converges weakly in* M to an element $u \in M$ (in symbols, $u_j \xrightarrow{w} u$ in M) if $\langle u_j, u^* \rangle \to \langle u, u^* \rangle$ in $\mathbb{K}$ as $j \to \infty$ for all $u^* \in M^*$; if this is the case then it is known that $\|u\| \le \liminf_{j\to\infty} \|u_j\|$.

Since a normed linear space $(M, \|\cdot\|)$ is a metric semigroup, the notions of the Vitali-type n-th variation, $|\alpha|$-th variation for $0 \ne \alpha \le 1$ and the total variation of a map $f : I_a^b \to M$ are introduced as above with respect to the induced metric $d(u,v) = \|u - v\|$, $u, v \in M$.

Theorem 2. *Suppose $(M, \|\cdot\|)$ is a reflexive separable Banach space with separable dual M^* and $\{f_j\}$ is a sequence of maps from I_a^b into M. If $\{f_j\}$ satisfies condition* (2.4) *from Theorem* 1 *and*

$$c(x) \equiv \sup_{j\in\mathbb{N}} \|f_j(x)\| \quad \text{is finite for all } x \in I_a^b, \tag{2.5}$$

then there exists a subsequence of $\{f_j\}$, again denoted by $\{f_j\}$, and a map $f \in \mathrm{BV}(I_a^b; M)$ satisfying $\mathrm{TV}(f, I_a^b) \le C$ such that

$$f_j(x) \xrightarrow{w} f(x) \quad \text{in } M \text{ for all } x \in I_a^b. \tag{2.6}$$

This theorem will be proved in Section 4. It is an extension of a weak selection principle from [6, Chapter 1, Theorem 3.5] for maps of bounded Jordan variation of one real variable. More comments and remarks on Theorems 1 and 2 can be found in Section 4.

3. Properties of mixed differences and the total variation

In this section we collect main ingredients of the proof of Theorem 1. These are relations between mixed differences of all orders and properties of the total variation (2.3). For real valued functions of n variables the main properties of mixed differences of all orders were elaborated in [1, 11, 17, 22, 28, 32, 38] and for metric semigroup valued maps of two variables—in [5, 13, 16, 35]. For our purposes we need their variants in the multiindex notation, as presented in [17] with $M = \mathbb{R}$, for maps of n variables with values in a metric semigroup.

First, we recall several definitions and results for real valued functions. A function $g : I_a^b \to \mathbb{R}$ is said to be *totally monotone* (cf., e.g., [17, Part II, Section 3], [32]) if, given $0 \ne \alpha \le 1$ and $x, y \in I_a^b$ with $x \le y$, we have:

$$(-1)^{|\alpha|} \sum_{0\le\theta\le\alpha} (-1)^{|\theta|} g\big(x + \theta(y - x)\big) \ge 0. \tag{3.1}$$

For real valued functions the sum in (3.1) (with no factor $(-1)^{|\alpha|}$) is called the $|\alpha|$*-th mixed difference* (in the sense of Vitali, Hardy and Krause) of g_α^x on the rectangle $I_x^y\lfloor\alpha$ and denoted by $\mathrm{md}_{|\alpha|}(g_\alpha^x, I_x^y\lfloor\alpha)$ (however, note the difference with (3.4) in the general case). In this case the *Vitali n-th variation* $V_n(g, I_a^b)$ of g on I_a^b is defined as in (2.2) with the mixed difference at the right-hand side of (2.2) replaced by $\big|\mathrm{md}_n(g, I_{x[\sigma-1]}^{x[\sigma]})\big|$. The other definitions related to the bounded variation context remain the same as above, and so, we keep the same notation for real valued functions as well.

Denote by $\mathrm{Mon}(I_a^b;\mathbb{R})$ the set of all totally monotone real valued functions on I_a^b. It is known (e.g., from the references above) that if $g\in\mathrm{Mon}(I_a^b;\mathbb{R})$, then $g\in\mathrm{BV}(I_a^b;\mathbb{R})$, the value at the left-hand side of (3.1) is equal to $V_{|\alpha|}(g_\alpha^x, I_x^y\lfloor\alpha)$, $g(x)\le g(y)$ and $\mathrm{TV}(g, I_x^y) = g(y)-g(x)$ for all $x, y\in I_a^b$ with $x\le y$.

The following Helly selection principle in the class $\mathrm{Mon}(I_a^b;\mathbb{R})$ is due to Leonov [32, Lemma 3] (for totally monotone functions of two variables it was established in [28, III.6.5] and [29, Theorem 3.1]):

Theorem A. *An infinite uniformly bounded family of totally monotone functions on I_a^b contains a sequence, which converges pointwise on I_a^b to a function from* $\mathrm{Mon}(I_a^b;\mathbb{R})$.

It was shown in [32, Corollary 2] that the linear space $\mathrm{BV}(I_a^b;\mathbb{R})$ equipped with the norm $\|g\| = |g(a)| + \mathrm{TV}(g, I_a^b)$, $g\in\mathrm{BV}(I_a^b;\mathbb{R})$, is a Banach space. This assertion was refined in [17, Part I, Theorem 1]: the space $\mathrm{BV}(I_a^b;\mathbb{R})$ is a Banach algebra with respect to the norm $\|\cdot\|$, and $\|g\cdot g'\|\le 2^n\|g\|\cdot\|g'\|$ for all $g, g'\in\mathrm{BV}(I_a^b;\mathbb{R})$.

Theorem A implies Helly's selection principle in the space $\mathrm{BV}(I_a^b;\mathbb{R})$ [32, Theorem 4]: *an infinite family of functions from* $\mathrm{BV}(I_a^b;\mathbb{R})$*, which is bounded under the norm* $\|\cdot\|$*, contains a pointwise convergent sequence, whose pointwise limit belongs to* $\mathrm{BV}(I_a^b;\mathbb{R})$. The crucial observation in the proof of this result is that, given $g\in\mathrm{BV}(I_a^b;\mathbb{R})$, if we set $\nu_g(x) = \mathrm{TV}(g, I_a^x)$ and $\pi_g(x) = \nu_g(x) - g(x)$, $x\in I_a^b$, then ([32, Theorem 3]) $\nu_g, \pi_g\in\mathrm{Mon}(I_a^b;\mathbb{R})$, and Jordan's decomposition holds: $g = \nu_g - \pi_g$ on I_a^b; then Theorem A applies to the uniformly bounded families of functions $\{\nu_g\}$ and $\{\pi_g\}$ in the standard way.

Now let us consider the case of maps $f: I_a^b\to M$ of finite total variation valued in a metric semigroup $(M, d, +)$. Clearly, there is no counterpart of Jordan's decomposition for these maps, and so, in order to prove Theorem 1,

we ought to argue in a completely different way. It will be seen later that, along with Theorem A, the following four Theorems B through E are the main ingredients in the proof of Theorem 1 (in a certain sense replacing the arguments involving Jordan's decomposition).

Theorem B. *If* $f \in \mathrm{BV}(I_a^b; M)$*,* $x, y \in I_a^b$ *and* $x \le y$*, then*

$$d(f(x), f(y)) \le \sum_{0 \ne \alpha \le 1} \mathrm{md}_{|\alpha|}(f_\alpha^x, I_x^y \lfloor \alpha) \le \mathrm{TV}(f, I_x^y).$$

This theorem will be proved in Section 5. It is a generalization of the well-known property of maps of bounded Jordan variation of one variable and a counterpart of Leonov's (in)equalities established in [32, Theorem 2 and Corollary 5] for real valued functions of n variables (cf. also [17, Part I, Lemma 6 and (3.5)]). The inequalities in Theorem B are also known for metric semigroup valued maps of two variables [5, 16]. However, in the general case Theorem B needs a different proof as compared to the cases of maps of one or two variable(s) or $M = \mathbb{R}$.

Theorem C. *If* $f \in \mathrm{BV}(I_a^b; M)$*,* $x, y \in I_a^b$*,* $x \le y$*, and* $0 \ne \gamma \le 1$*, then*

$$\begin{aligned} \sum_{0 \ne \alpha \le \gamma} V_{|\alpha|}(f_\alpha^x, I_x^y \lfloor \alpha) &= \mathrm{TV}\big(f, I_x^{x+\gamma(y-x)}\big) \\ &\le \mathrm{TV}(f, I_a^{x+\gamma(y-x)}) - \mathrm{TV}(f, I_a^x). \end{aligned} \tag{3.2}$$

Theorem D. *If* $f \in \mathrm{BV}(I_a^b; M)$ *and if we set* $\nu_f(x) = \mathrm{TV}(f, I_a^x)$*,* $x \in I_a^b$*, then for* $\nu_f : I_a^b \to \mathbb{R}$*, called the total variation function of* f*, we have:* $\nu_f \in \mathrm{Mon}(I_a^b; \mathbb{R})$ *and* $\mathrm{TV}(\nu_f, I_a^b) = \mathrm{TV}(f, I_a^b)$*.*

These two theorems are extensions of two more properties of the Jordan variation for maps of one variable; in this case (3.2) is actually the equality known as the *additivity* of Jordan's variation (e.g., [37, Theorem 83_2]). On the other hand, Theorem C is a counterpart of Chistyakov's inequality [17, Part II, Lemma 8] and Theorem D is a generalization of Theorem 3 from [32] and Corollary 11 from [17, Part II] given for $M = \mathbb{R}$. For metric semigroup valued maps of two variables cf. [5, inequalities (11), (13) and Theorem 1].

The proof of Theorem C is identical with the proof of Lemma 8 from [17, Part II] and the proof of Theorem D is identical with the proofs of Lemma 9 and Corollaries 10 and 11 from [17, Part II] when $M = \mathbb{R}$, and so, they are

omitted. However, it is to be noted that these proofs rely on (1) equality (3.2) from [17, Part I, Lemma 5], (2) Lemma 7 from [17, Part I], and (3) the well-known property of the *additivity* of $|\alpha|$-th variation $V_{|\alpha|}$ for each $0 \neq \alpha \leq 1$ for real valued functions of n variables. For metric semigroup valued maps assertions (1), (2) and (3) need a proper interpretation and different, more subtle and hard proofs. Their respective counterparts are presented below as Lemmas 1, 2 and 3.

In the first lemma and throughout the paper we use the following short notations: given $0 \neq \alpha \leq 1$, the sum over '$\mathrm{ev}\,\theta \leq \alpha$' denotes the sum over '$\theta \in \mathcal{E}(n)$ s.t. $\theta \leq \alpha$', where 's.t.' is the usual abbreviation for 'such that', and a similar convention applies to the sum over '$\mathrm{od}\,\theta \leq \alpha$'.

Lemma 1. *If $f : I_a^b \to M$, $x, y \in I_a^b$, $x \leq y$, $z \in I_a^b$ and $0 \neq \alpha \leq 1$, then*

$$\begin{aligned} \mathrm{md}_{|\alpha|}(f_\alpha^z, I_x^y \lfloor \alpha) \;=\; d\biggl(& \sum_{\mathrm{ev}\,\theta \leq \alpha} f\bigl(z + \alpha(x - z) + \theta(y - x)\bigr), \\ & \sum_{\mathrm{od}\,\theta \leq \alpha} f\bigl(z + \alpha(x - z) + \theta(y - x)\bigr)\biggr). \end{aligned} \tag{3.3}$$

In particular, if $z = a$ or $z = x$, we have, respectively,

$$\mathrm{md}_{|\alpha|}(f_\alpha^a, I_x^y \lfloor \alpha) = \mathrm{md}_{|\alpha|}(f_\alpha^{a+\alpha(x-a)}, I_{a+\alpha(x-a)}^y \lfloor \alpha),$$

$$\mathrm{md}_{|\alpha|}(f_\alpha^x, I_x^y \lfloor \alpha) = d\biggl(\sum_{\mathrm{ev}\,\theta \leq \alpha} f\bigl(x + \theta(y - x)\bigr), \sum_{\mathrm{od}\,\theta \leq \alpha} f\bigl(x + \theta(y - x)\bigr)\biggr). \tag{3.4}$$

The proof of Lemma 1 is the same as in [17, Part I, Lemma 5] (details are omitted): we have to note only that $\theta' \in \mathbb{N}_0^{|\alpha|}$ and $|\theta'|$ is even (odd) if and only if there exists a unique $\theta \in \mathbb{N}_0^n$ s.t. $\theta \leq \alpha$, $|\theta|$ is even (odd, respectively) and $\theta' = \theta \lfloor \alpha$, and apply definition (2.1) where n is replaced by $|\alpha|$.

Since the total variation (2.3) is defined via truncated maps with the base at the point a, in our next lemma we present a counterpart of Chistyakov's equality [17, Part I, Lemma 7] exhibiting the relation between the mixed difference $\mathrm{md}_{|\alpha|}(f_\alpha^x, I_x^y \lfloor \alpha)$ and certain mixed differences of maps f_β^a with the base at a for some $0 \neq \beta \leq 1$.

Lemma 2. *If $f : I_a^b \to M$, $0 \neq \alpha \leq 1$ and $x, y \in I_a^b$ with $x \leq y$, then*

$$\mathrm{md}_{|\alpha|}(f_\alpha^x, I_x^y \lfloor \alpha) \leq \sum_{\alpha \leq \beta \leq 1} \mathrm{md}_{|\beta|}\bigl(f_\beta^a, I_{a+\alpha(x-a)}^{x+\alpha(y-x)} \lfloor \beta\bigr).$$

The proof of Lemma 2 is postponed until Section 6.

The *additivity* property of $|\alpha|$-th variation $V_{|\alpha|}$ for each $0 \neq \alpha \leq 1$, to be proved in Section 7, is expressed in the following

Lemma 3. *Given $f : I_a^b \to M$, $x, y \in I_a^b$ with $x < y$, $z \in I_a^b$ and $0 \neq \alpha \leq 1$, if $\{x[\sigma]\}_{\sigma=0}^{\kappa}$ is a net partition of I_x^y, then*

$$V_{|\alpha|}(f_\alpha^z, I_x^y \lfloor \alpha) = \sum_{1\lfloor\alpha \leq \sigma\lfloor\alpha \leq \kappa\lfloor\alpha} V_{|\alpha|}(f_\alpha^z, I_{x[\sigma-1]}^{x[\sigma]} \lfloor \alpha), \tag{3.5}$$

where the summation is taken only over those σ_i in the range $1 \leq \sigma_i \leq \kappa_i$ with $i \in \{1, \dots, n\}$, for which $\alpha_i = 1$.

The final ingredient in the proof of Theorem 1 is the *sequential lower semicontinuity* of the total variation $\mathrm{TV}(\cdot, I_a^b)$ to be established in Section 8:

Theorem E. *If a sequence of maps $\{f_j\}$ from I_a^b into M converges pointwise on I_a^b to a map $f : I_a^b \to M$, then $\mathrm{TV}(f, I_a^b) \leq \liminf_{j\to\infty} \mathrm{TV}(f_j, I_a^b)$.*

Now we are in a position to prove Theorems 1 and 2.

4. Proofs of Theorems 1 and 2

Proof of Theorem 1. We divide the proof into four steps for clarity.

1. We apply the induction argument on the dimension n of the basic rectangle $I_a^b \subset \mathbb{R}^n$. For $n = 1$ Theorem 1 was established in [10, Theorem 5.1] (and refined in [7, Theorem 1] and [15, Theorem 1.3]) in the case when (M, d) is an arbitrary metric space, and for $n = 2$ it was proved in [5, Theorem 2]. Now, suppose that $n \geq 3$ and Theorem 1 is already established for domain rectangles of dimension $\leq n - 1$.

Given $j \in \mathbb{N}$, we let ν_j be the total variation function of f_j on I_a^b, i.e., $\nu_j(x) = \mathrm{TV}(f_j, I_a^x)$ for all $x \in I_a^b$. By Theorem D and condition (2.4), the sequence $\{\nu_j\} \subset \mathrm{Mon}(I_a^b; \mathbb{R})$ is uniformly bounded (by C), and so, by Theorem A, there exist a subsequence of $\{\nu_j\}$ and the corresponding subsequence of $\{f_j\}$, again denoted as the whole sequences $\{\nu_j\}$ and $\{f_j\}$, respectively, and a function $\nu \in \mathrm{Mon}(I_a^b; \mathbb{R})$ s.t.

$$\lim_{j\to\infty} \nu_j(x) = \nu(x) \quad \text{for all} \quad x \in I_a^b. \tag{4.1}$$

It is known ([4], [28, III.5.4], [40]) that the set of discontinuity points of any totally monotone function on $I_a^b \subset \mathbb{R}^n$ lies on at most a countable set

of hyperplanes of dimension $n-1$ parallel to the coordinate axes. Given $i \in \{1, \dots, n\}$, denote by Z_i the union of the set of all rational points of the interval $[a_i, b_i]$, the two-point set $\{a_i, b_i\}$ and the set of those points $z_i \in [a_i, b_i]$, for which the hyperplane

$$H_i(z_i) = [a_1, b_1] \times \cdots \times [a_{i-1}, b_{i-1}] \times \{z_i\} \times [a_{i+1}, b_{i+1}] \times \cdots \times [a_n, b_n] \quad (4.2)$$

contains points of discontinuity of ν. Clearly, the sets $Z_i \subset [a_i, b_i]$ are countable and dense in $[a_i, b_i]$, and so, we may assume that $Z_i = \{z_i(k)\}_{k=1}^{\infty}$.

2. In order to apply the induction hypothesis, we need an estimate on the $(n-1)$-dimensional total variation of any function $f = f_j$ from the sequence $\{f_j\}$ 'over the hyperplane' (4.2) in the sense to be made precise below. This is done as follows.

Let us fix $i \in \{1, \dots, n\}$ and set $1^i = (1, \dots, 1, 0, 1, \dots, 1)$, where 0 is the i-th coordinate of 1^i and the other coordinates of 1^i are equal to 1. Note that $|1^i| = n-1$. Given $z_i \in Z_i$, we put

$$\overline{a} \equiv \overline{a}(z_i) = (a_1, \dots, a_{i-1}, z_i, a_{i+1}, \dots, a_n). \quad (4.3)$$

The map $f_{1^i}^{\overline{a}} : I_a^b \lfloor 1^i \to M$ with the base at $\overline{a}$, truncated by 1^i, is defined on the $(n-1)$-dimensional rectangle $I_a^b \lfloor 1^i \subset \mathbb{R}^{n-1}$ and given by: if $x \in I_a^b$, then $x \lfloor 1^i \in I_a^b \lfloor 1^i$ and

$$f_{1^i}^{\overline{a}}(x \lfloor 1^i) = f(\overline{a} + 1^i(x - \overline{a})) = f(x_1, \dots, x_{i-1}, z_i, x_{i+1}, \dots, x_n). \quad (4.4)$$

The $(n-1)$-dimensional total variation of $f_{1^i}^{\overline{a}}$ on $I_a^b \lfloor 1^i$ is equal to

$$\mathrm{TV}_{n-1}(f_{1^i}^{\overline{a}}, I_a^b \lfloor 1^i) = \sum_{0 \neq \alpha \leq 1} V_{|\alpha|}\big((f_{1^i}^{\overline{a}})_\alpha^{a \lfloor 1^i}, (I_a^b \lfloor 1^i) \lfloor \alpha\big), \quad (4.5)$$

where the summation is taken over $(n-1)$-dimensional multiindices $\alpha = (\alpha_1, \dots, \alpha_{n-1})$ s.t. $0_{n-1} \neq \alpha \leq 1_{n-1}$, i.e., $\alpha \in \mathcal{A}(n-1)$ (this is the only instance and exception when the summation is over $(n-1)$-dimensional multiindices). Given $\alpha \in \mathcal{A}(n-1)$, we set $\overline{\alpha} = (\alpha_1, \dots, \alpha_{i-1}, 0, \alpha_i, \dots, \alpha_{n-1})$, where 0 occupies the i-th place, and note that $\alpha = \overline{\alpha} \lfloor 1^i$. We have

$$(f_{1^i}^{\overline{a}})_\alpha^{a \lfloor 1^i} = f_{\overline{\alpha}}^{\overline{a}} \quad \text{on} \quad (I_a^b \lfloor 1^i) \lfloor \alpha = I_a^b \lfloor \overline{\alpha} = I_{\overline{a}}^b \lfloor \overline{\alpha}.$$

In fact, given $x \in I_a^b$, we find $x\lfloor\overline{\alpha} = (x\lfloor 1^i)\lfloor\alpha$ and

$$\begin{aligned}(f_{1^i}^{\overline{a}})_\alpha^{a\lfloor 1^i}(x\lfloor\overline{\alpha}) &= (f_{1^i}^{\overline{a}})_{\overline{\alpha}\lfloor 1^i}^{a\lfloor 1^i}((x\lfloor 1^i)\lfloor\alpha) \\ &= f_{1^i}^{\overline{a}}\big((a\lfloor 1^i) + (\overline{\alpha}\lfloor 1^i)[(x\lfloor 1^i) - (a\lfloor 1^i)]\big) \\ &= f_{1^i}^{\overline{a}}\big([a + \overline{\alpha}\,(x-a)]\lfloor 1^i\big) \\ &= f\big(\overline{a} + 1^i[a + \overline{\alpha}\,(x-a) - \overline{a}]\big). \end{aligned} \tag{4.6}$$

Since $\overline{a} + 1^i(a - \overline{a}) = \overline{a}$ and $1^i\,\overline{\alpha} = \overline{\alpha}$, we get

$$\begin{aligned}\overline{a} + 1^i[a + \overline{\alpha}\,(x-a) - \overline{a}] &= \overline{a} + 1^i(a - \overline{a}) + 1^i\,\overline{\alpha}\,(x-a) \\ &= \overline{a} + \overline{\alpha}\,(x-a) = \overline{a} + \overline{\alpha}\,(x - \overline{a}),\end{aligned}$$

and so, the value (4.6) is equal to

$$f(\overline{a} + \overline{\alpha}\,(x - \overline{a})) = f_{\overline{\alpha}}^{\overline{a}}(x\lfloor\overline{\alpha}).$$

It follows that the $|\alpha|$-th variation at the right-hand side of (4.5) is equal to

$$V_{|\alpha|}\big((f_{1^i}^{\overline{a}})_\alpha^{a\lfloor 1^i}, (I_a^b\lfloor 1^i)\lfloor\alpha\big) = V_{|\overline{\alpha}|}(f_{\overline{\alpha}}^{\overline{a}}, I_{\overline{a}}^b\lfloor\overline{\alpha}).$$

Noting that the set $\mathcal{A}(n-1)$ is bijective to the set of those $\overline{\alpha} \in \mathcal{A}(n)$, for which $0 \neq \overline{\alpha} \leq 1^i$, and applying Theorem C with $x = \overline{a}$, $y = b$ and $\gamma = 1^i$, we get:

$$\begin{aligned}\mathrm{TV}_{n-1}(f_{1^i}^{\overline{a}}, I_a^b\lfloor 1^i) &= \sum_{0\neq\overline{\alpha}\leq 1^i} V_{|\overline{\alpha}|}(f_{\overline{\alpha}}^{\overline{a}}, I_{\overline{a}}^b\lfloor\overline{\alpha}) = \mathrm{TV}(f, I_{\overline{a}}^{\overline{a}+1^i(b-\overline{a})}) \\ &\leq \mathrm{TV}(f, I_a^{\overline{a}+1^i(b-\overline{a})}) - \mathrm{TV}(f, I_a^{\overline{a}}) \leq \mathrm{TV}(f, I_a^b).\end{aligned} \tag{4.7}$$

Thus, given $j \in \mathbb{N}$ and $i \in \{1, \ldots, n\}$, setting back $f = f_j$, by virtue of (4.3), (4.7) and (2.4), we find, for all $z_i \in Z_i$ and $\overline{a} = \overline{a}(z_i)$:

$$\mathrm{TV}_{n-1}\big((f_j)_{1^i}^{\overline{a}(z_i)}, I_a^b\lfloor 1^i\big) \leq C < \infty. \tag{4.8}$$

3. Now, we make use of the diagonal processes. For $i = 1$ and $z_1 = z_i(1) = z_1(1) \in Z_1$ the sequence $\{(f_j)_{1^i}^{\overline{a}(z_i(1))}\}_{j=1}^\infty = \{(f_j)_{1^1}^{\overline{a}(z_1(1))}\}_{j=1}^\infty$ satisfies the uniform estimate (4.8) on the rectangle $I_a^b\lfloor 1^1$ of dimension $n-1$ and, since each map from this sequence is of the form (4.4) with $z_i = z_1 = z_1(1)$, then it follows from the assumptions of Theorem 1 that the sequence under consideration is pointwise precompact on $I_a^b\lfloor 1^1$. By the induction hypothesis, the

sequence $\{f_j\}$ contains a subsequence, denoted by $\{f_j^1\}$, s.t. $(f_j^1)_{1^1}^{\overline{a}(z_1(1))}$ converges pointwise on $I_a^b\lfloor 1^1$ to a map from $I_a^b\lfloor 1^1$ into M of $(n-1)$-dimensional finite total variation on $I_a^b\lfloor 1^1$. Since, by (4.4),

$$(f_j^1)_{1^1}^{\overline{a}(z_1(1))}(x_2,\dots,x_n) = (f_j^1)_{1^1}^{\overline{a}(z_1(1))}(x\lfloor 1^1) = f_j^1(z_1(1),x_2,\dots,x_n)$$

with $x=(x_1,\dots,x_n)\in I_a^b$ and $x_i\in[a_i,b_i]$ for $i\in\{2,\dots,n\}$, then the pointwise convergence above means, actually, that the sequence $\{f_j^1\}$ converges pointwise on the hyperplane $H_1(z_1(1))=\{z_1(1)\}\times[a_2,b_2]\times\cdots\times[a_n,b_n]$.

Inductively, if $k\geq 2$ and a subsequence $\{f_j^{k-1}\}_{j=1}^{\infty}$ of $\{f_j\}$, which is pointwise convergent on $\bigcup_{l=1}^{k-1} H_1(z_1(l))$, is already chosen, then the sequence $\{(f_j^{k-1})_{1^1}^{\overline{a}(z_1(k))}\}_{j=1}^{\infty}$ satisfies the uniform estimate (4.8) on the rectangle $I_a^b\lfloor 1^1$, where f_j is replaced by f_j^{k-1} and $\overline{a}(z_i)$—by $\overline{a}(z_1(k))$. Moreover, since, as above, the sequence is pointwise precompact on $I_a^b\lfloor 1^1$, then, by the induction hypothesis, there exists a subsequence $\{f_j^k\}_{j=1}^{\infty}$ of $\{f_j^{k-1}\}_{j=1}^{\infty}$ s.t. $(f_j^k)_{1^1}^{\overline{a}(z_1(k))}$ converges pointwise on $I_a^b\lfloor 1^1$ as $j\to\infty$ to a map from $I_a^b\lfloor 1^1$ into M of $(n-1)$-dimensional finite total variation on $I_a^b\lfloor 1^1$. Again, as above, this pointwise convergence means that the sequence $\{f_j^k\}_{j=1}^{\infty}$ converges pointwise on the hyperplane $H_1(z_1(k))$ and, as a consequence, on the set $\bigcup_{l=1}^{k} H_1(z_1(l))$ as well. We infer that the diagonal sequence $\{f_j^j\}_{j=1}^{\infty}$, which is a subsequence of the original sequence $\{f_j\}$, converges pointwise on the set $H_1(Z_1)=\bigcup_{z_1\in Z_1} H_1(z_1)=\bigcup_{l=1}^{\infty} H_1(z_1(l))$; in fact, given $(z_1,x_2,\dots,x_n)\in H_1(Z_1)$, we have $z_1=z_1(k)\in Z_1$ for some $k\in\mathbb{N}$ and $(x_2,\dots,x_n)\in I_a^b\lfloor 1^1$, and so, noting that $\{f_j^j\}_{j=k}^{\infty}$ is a subsequence of $\{f_j^k\}_{j=1}^{\infty}$, we find that

$$f_j^j(z_1,x_2,\dots,x_n) = (f_j^j)_{1^1}^{\overline{a}(z_1(k))}(x_2,\dots,x_n)$$

converges in M as $j\to\infty$.

Let us denote the diagonal sequence $\{f_j^j\}_{j=1}^{\infty}$ extracted in the last paragraph again by $\{f_j\}$. Then we let $i=2$, $z_2=z_i(1)=z_2(1)\in Z_2$ and, beginning with the sequence $\{(f_j)_{1^i}^{\overline{a}(z_i(1))}\}_{j=1}^{\infty}=\{(f_j)_{1^2}^{\overline{a}(z_2(1))}\}_{j=1}^{\infty}$, apply the above arguments of this step. Doing this, we will end up with a diagonal sequence, a subsequence of the original sequence $\{f_j\}$, again denoted by $\{f_j\}$, which converges pointwise on $H_1(Z_1)\cup H_2(Z_2)$. Now suppose that for some $i\in\{2,\dots,n-1\}$ we have already extracted a (diagonal) subsequence of $\{f_j\}$, again denoted by $\{f_j\}$, which converges pointwise on the set $H_1(Z_1)\cup\cdots\cup H_{i-1}(Z_{i-1})$. Then we let $z_i=z_i(1)\in Z_i$ and apply the above

arguments of this step to the sequence $\{(f_j)_{1^i}^{\overline{a}(z_i(1))}\}_{j=1}^{\infty}$: a subsequence of the original sequence $\{f_j\}$ converges pointwise on the set $H_1(Z_1)\cup\dots\cup H_i(Z_i)$. In this way after finitely many steps we obtain a subsequence of the original sequence $\{f_j\}$, again denoted by $\{f_j\}$, which converges pointwise on the set $H(Z)=\bigcup_{i=1}^n H_i(Z_i)$.

4. Finally, let us show that the sequence $\{f_j\}$ from the end of Step 3 converges at each point $y\in I_a^b\setminus H(Z)$. Note that y is a point of continuity of the function ν from (4.1) s.t. its coordinates $a_i<y_i<b_i$ are irrational for all $i\in\{1,\dots,n\}$. Due to the density of $H(Z)$ in I_a^b, the continuity of ν at y and properties of totally monotone functions, given $\varepsilon>0$, there exists $x=x(\varepsilon)\in H(Z)$ with $x<y$ s.t. $0\le\nu(y)-\nu(x)\le\varepsilon$. By virtue of (4.1), choose a number $j_0(\varepsilon)\in\mathbb{N}$ s.t. $|\nu_j(y)-\nu(y)|\le\varepsilon$ and $|\nu(x)-\nu_j(x)|\le\varepsilon$ for all $j\ge j_0(\varepsilon)$. By Theorems B and C with $\gamma=1$, for all $j\ge j_0(\varepsilon)$ we have:

$$\begin{aligned}d(f_j(x),f_j(y)) &\le \mathrm{TV}(f_j,I_x^y)\le\mathrm{TV}(f_j,I_a^y)-\mathrm{TV}(f_j,I_a^x)=\nu_j(y)-\nu_j(x)\\ &\le |\nu_j(y)-\nu(y)|+(\nu(y)-\nu(x))+|\nu(x)-\nu_j(x)|\le 3\varepsilon.\end{aligned}$$

Since $x\in H(Z)$ and, as it was shown in Step 3, the sequence $\{f_j(x)\}_{j=1}^{\infty}$ is convergent in M, it is Cauchy, and so, there exists a number $j_1(\varepsilon)\in\mathbb{N}$ s.t. $d(f_j(x),f_{j'}(x))\le\varepsilon$ for all $j\ge j_1(\varepsilon)$ and $j'\ge j_1(\varepsilon)$. It follows that if $J(\varepsilon)=\max\{j_0(\varepsilon),j_1(\varepsilon)\}$, $j\ge J(\varepsilon)$ and $j'\ge J(\varepsilon)$, then we have:

$$\begin{aligned}d(f_j(y),f_{j'}(y)) &\le d(f_j(y),f_j(x))+d(f_j(x),f_{j'}(x))+d(f_{j'}(x),f_{j'}(y))\\ &\le 3\varepsilon+\varepsilon+3\varepsilon=7\varepsilon.\end{aligned}$$

Thus, the sequence $\{f_j(y)\}_{j=1}^{\infty}$ is Cauchy in the metric space M, and so, since it is also precompact by the assumption, it is convergent in M.

It follows from here and the end of Step 3 that the sequence $\{f_j(y)\}_{j=1}^{\infty}$ converges in M at each point $y\in(I_a^b\setminus H(Z))\cup H(Z)=I_a^b$, i.e., the sequence $\{f_j\}$, which is a subsequence of the original sequence $\{f_j\}$, converges pointwise on I_a^b. Let us denote the pointwise limit of $\{f_j\}$ by $f:I_a^b\to M$. Then, by virtue of Theorem E and assumption (2.4), we find

$$\mathrm{TV}(f,I_a^b)\le\liminf_{j\to\infty}\mathrm{TV}(f_j,I_a^b)\le C,$$

and so, $f\in\mathrm{BV}(I_a^b;M)$.

This completes the proof of Theorem 1. $\square$

Remark 4.1. In Theorem 1 the precompactness of the sets $\{f_j(x)\}_{j=1}^{\infty}$ at all points $x \in I_a^b$ cannot be replaced by the closedness and boundedness even at a single point of I_a^b. The corresponding examples for maps of one variable are constructed in [7, Section 3], [10, Section 5] and [15, Section 1] and can be easily adapted for maps of several variables.

Proof of Theorem 2. The proof is adapted for the situation under consideration from the proof of Theorem 7 from [18].

1. In this step we show that, given $j \in \mathbb{N}$ and $u^* \in M^*$, we have:

$$\mathrm{TV}(\langle f_j(\cdot), u^*\rangle, I_a^b) \le \mathrm{TV}(f_j, I_a^b)\|u^*\| \le C\|u^*\|, \tag{4.9}$$

where the function $\langle f_j(\cdot), u^*\rangle : I_a^b \to \mathbb{K}$ is given by $\langle f_j(\cdot), u^*\rangle(x) = \langle f_j(x), u^*\rangle$, $x \in I_a^b$, and C is the constant from (2.4).

In fact, given $0 \ne \alpha \le 1$ and $x, y \in I_a^b$ with $x < y$, by virtue of (3.3) where $d(u, v)$ is replaced by the absolute value $|u - v|$ in $\mathbb{K}$ and later on—by the norm in M, we get:

$$\begin{aligned}
\mathrm{md}_{|\alpha|}(\langle f_j(\cdot), u^*\rangle_\alpha^a, I_x^y\lfloor\alpha) &= \Big|\sum_{0\le\theta\le\alpha}(-1)^{|\theta|}\langle f_j(a+\alpha\,(x-a)+\theta(y-x)), u^*\rangle\Big| \\
&= \Big|\Big\langle \sum_{0\le\theta\le\alpha}(-1)^{|\theta|} f_j(a+\alpha\,(x-a)+\theta(y-x)), u^*\Big\rangle\Big| \\
&\le \Big\|\sum_{0\le\theta\le\alpha}(-1)^{|\theta|} f_j(a+\alpha\,(x-a)+\theta(y-x))\Big\| \cdot \|u^*\| \\
&= \mathrm{md}_{|\alpha|}((f_j)_\alpha^a, I_x^y\lfloor\alpha)\|u^*\|.
\end{aligned}$$

It follows that if $\mathcal{P} = \{x[\sigma]\}_{\sigma=0}^{\kappa}$ is a net partition of I_a^b, then $\mathcal{P}\lfloor\alpha = \{x[\sigma]\lfloor\alpha\}_{\sigma\lfloor\alpha=0}^{\kappa\lfloor\alpha}$ is a net partition of $I_a^b\lfloor\alpha$, and so, setting $x = x[\sigma-1]$ and $y = x[\sigma]$ in the calculations above, we find

$$\begin{aligned}
\sum_{1\lfloor\alpha\le\sigma\lfloor\alpha\le\kappa\lfloor\alpha} \mathrm{md}_{|\alpha|}\big(\langle f_j(\cdot), u^*\rangle_\alpha^a, I_{x[\sigma-1]}^{x[\sigma]}\lfloor\alpha\big) &\le \sum_{1\lfloor\alpha\le\sigma\lfloor\alpha\le\kappa\lfloor\alpha} \mathrm{md}_{|\alpha|}\big((f_j)_\alpha^a, I_{x[\sigma-1]}^{x[\sigma]}\lfloor\alpha\big)\|u^*\| \\
&\le V_{|\alpha|}((f_j)_\alpha^a, I_a^b\lfloor\alpha)\|u^*\|,
\end{aligned}$$

the summation over $\sigma\lfloor\alpha$ being taken only over those coordinates σ_i in the range $1 \le \sigma_i \le \kappa_i$ with $i \in \{1, \dots, n\}$, for which $\alpha_i = 1$. Since $\mathcal{P}$ is an arbitrary partition of I_a^b, we get:

$$V_{|\alpha|}\big(\langle f_j(\cdot), u^*\rangle_\alpha^a, I_a^b\lfloor\alpha\big) \le V_{|\alpha|}\big((f_j)_\alpha^a, I_a^b\lfloor\alpha\big)\|u^*\|,$$

and so, inequality (4.9) follows from the definition of the total variation.

Moreover, by virtue of (2.5), we have:

$$|\langle f_j(x), u^*\rangle| \le \|f_j(x)\| \cdot \|u^*\| \le c(x)\|u^*\|, \qquad x \in I_a^b, \quad u^* \in M^*, \tag{4.10}$$

and so, the sequence $\{\langle f_j(\cdot), u^*\rangle\}_{j=1}^\infty$ of functions from I_a^b into (metric semigroup) $\mathbb{K}$ is pointwise bounded on I_a^b and, hence, pointwise precompact for each $u^* \in M^*$.

Taking this and (4.9) into account and applying Theorem 1 to the sequence $\{\langle f_j(\cdot), u^*\rangle\}_{j=1}^\infty$ for any given $u^* \in M^*$, we extract a subsequence of $\{f_j\}$, denoted by $\{f_{j,u^*}\}$ (which depends on u^* in general), and find a function $g_{u^*} \in \mathrm{BV}(I_a^b; \mathbb{K})$ satisfying $\mathrm{TV}(g_{u^*}, I_a^b) \le C\|u^*\|$ s.t. $\langle f_{j,u^*}(x), u^*\rangle \to g_{u^*}(x)$ in $\mathbb{K}$ as $j \to \infty$ for all $x \in I_a^b$.

2. Making use of the diagonal process and the separability of M^*, let us get rid of the dependence of $\{f_{j,u^*}\}$ on $u^* \in M^*$. Let $\{u_k^*\}_{k=1}^\infty$ be a countable dense subset of M^*. By Step 1, for $u^* = u_1^*$ we get a subsequence $\{f_j^{(1)}\} = \{f_{j,u_1^*}\}_{j=1}^\infty$ of the original sequence $\{f_j\}$ and a function $g_{u_1^*} \in \mathrm{BV}(I_a^b; \mathbb{K})$ satisfying $\mathrm{TV}(g_{u_1^*}, I_a^b) \le C\|u_1^*\|$ s.t. $\langle f_j^{(1)}(x), u_1^*\rangle \to g_{u_1^*}(x)$ in $\mathbb{K}$ for all $x \in I_a^b$. Inductively, if $k \ge 2$ and a subsequence $\{f_j^{(k-1)}\}_{j=1}^\infty$ of $\{f_j\}$ is already chosen, then by virtue of (4.9) and (4.10), we have:

$$\mathrm{TV}(\langle f_j^{(k-1)}(\cdot), u_k^*\rangle, I_a^b) \le C\|u_k^*\|$$

and $|\langle f_j^{(k-1)}(x), u_k^*\rangle| \le c(x)\|u_k^*\|$, $x \in I_a^b$, for all $j \in \mathbb{N}$. By Theorem 1, applied to the sequence $\{\langle f_j^{(k-1)}(\cdot), u_k^*\rangle\}_{j=1}^\infty$, there exist a subsequence $\{f_j^{(k)}\}_{j=1}^\infty$ of $\{f_j^{(k-1)}\}_{j=1}^\infty$ and a function $g_{u_k^*} \in \mathrm{BV}(I_a^b; \mathbb{K})$ satisfying $\mathrm{TV}(g_{u_k^*}, I_a^b) \le C\|u_k^*\|$ s.t. $\langle f_j^{(k)}(x), u_k^*\rangle \to g_{u_k^*}(x)$ in $\mathbb{K}$ as $j \to \infty$ for all $x \in I_a^b$. Then the diagonal sequence $\{f_j^{(j)}\}_{j=1}^\infty$, again denoted by $\{f_j\}$, is a subsequence of the original sequence and satisfies the condition:

$$\langle f_j(x), u_k^*\rangle \to g_{u_k^*}(x) \quad \text{as } j \to \infty \text{ for all } x \in I_a^b \text{ and } k \in \mathbb{N}. \tag{4.11}$$

3. Now, given $u^* \in M^*$ and $x \in I_a^b$, let us show that the sequence $\{\langle f_j(x), u^*\rangle\}_{j=1}^\infty$ is Cauchy in $\mathbb{K}$. Taking into account (4.11) we may assume that $u^* \ne u_k^*$ for all $k \in \mathbb{N}$. Let $\varepsilon > 0$ be arbitrary. By the density of $\{u_k^*\}_{k=1}^\infty$ in M^*, there exists $k = k(\varepsilon) \in \mathbb{N}$ s.t. $\|u^* - u_k^*\| \le \varepsilon/(1 + 4c(x))$. By (4.11),

there exists $j_0 = j_0(\varepsilon) \in \mathbb{N}$ s.t. $|\langle f_j(x), u_k^*\rangle - \langle f_{j'}(x), u_k^*\rangle| \le \varepsilon/2$ for all $j \ge j_0$ and $j' \ge j_0$. It follows that for such j and j' we have:

$$\begin{aligned}
|\langle f_j(x), u^*\rangle - \langle f_{j'}(x), u^*\rangle| &\le |\langle f_j(x) - f_{j'}(x), u^* - u_k^*\rangle| \\
&\quad +|\langle f_j(x), u_k^*\rangle - \langle f_{j'}(x), u_k^*\rangle| \\
&\le \|f_j(x) - f_{j'}(x)\| \cdot \|u^* - u_k^*\| + \frac{\varepsilon}{2} \\
&\le 2c(x)\frac{\varepsilon}{1+4c(x)} + \frac{\varepsilon}{2} \le \varepsilon.
\end{aligned}$$

Thus, $\{\langle f_j(x), u^*\rangle\}_{j=1}^{\infty}$ is Cauchy in $\mathbb{K}$ and, hence, there exists an element of $\mathbb{K}$ denoted by $g_{u^*}(x)$ s.t. $\langle f_j(x), u^*\rangle \to g_{u^*}(x)$ in $\mathbb{K}$ as $j \to \infty$. In other words, we have shown that for each $u^* \in M^*$ there exists a function $g_{u^*} : I_a^b \to \mathbb{K}$ satisfying (cf. Theorem E and (4.9))

$$\mathrm{TV}(g_{u^*}, I_a^b) \le \liminf_{j\to\infty} \mathrm{TV}(\langle f_j(\cdot), u^*\rangle, I_a^b) \le C\|u^*\|$$

(and so, $g_{u^*} \in \mathrm{BV}(I_a^b; \mathbb{K})$) and

$$\lim_{j\to\infty} \langle f_j(x), u^*\rangle = g_{u^*}(x) \quad \text{in } \mathbb{K} \text{ for all } x \in I_a^b \text{ and } u^* \in M^*. \tag{4.12}$$

4. Let us prove (2.6), i.e., $f_j(x)$ converges weakly in M as $j \to \infty$ for all $x \in I_a^b$. By the reflexivity of M, we have $f_j(x) \in M = M^{**} \equiv \mathrm{L}(M^*; \mathbb{K})$ for all $j \in \mathbb{N}$. Define the functional $G_x : M^* \to \mathbb{K}$ by $G_x(u^*) = g_{u^*}(x)$ for all $u^* \in M^*$. By virtue of (4.12), we get

$$\lim_{j\to\infty} \langle f_j(x), u^*\rangle = g_{u^*}(x) = G_x(u^*) \quad \text{for all} \quad u^* \in M^*,$$

i.e., the sequence $\{f_j(x)\}_{j=1}^{\infty} \subset \mathrm{L}(M^*; \mathbb{K})$ converges pointwise on M^* to the operator $G_x : M^* \to \mathbb{K}$. By the Banach-Steinhaus uniform boundedness principle, $G_x \in \mathrm{L}(M^*; \mathbb{K}) = M$ and $\|G_x\| \le \liminf_{j\to\infty} \|f_j(x)\|$. Setting $f(x) = G_x$ for all $x \in I_a^b$, we find that $f : I_a^b \to M$ and

$$\lim_{j\to\infty} \langle f_j(x), u^*\rangle = G_x(u^*) = \langle G_x, u^*\rangle = \langle f(x), u^*\rangle \quad \text{in} \quad \mathbb{K} \tag{4.13}$$

for all $u^* \in M^*$ and $x \in I_a^b$, and so, $f_j(x) \xrightarrow{w} f(x)$ in M as $j \to \infty$ for all $x \in I_a^b$, which proves (2.6).

5. It remains to show that $f \in \mathrm{BV}(I_a^b; M)$ and $\mathrm{TV}(f, I_a^b) \le C$. By (4.13), we have: if $x, y \in I_a^b$ with $x < y$ and $0 \ne \alpha \le 1$, then

$$\sum_{0\le\theta\le\alpha} (-1)^{|\theta|} f_j\big(a+\alpha\,(x-a)+\theta(y-x)\big) \xrightarrow{w} \sum_{0\le\theta\le\alpha} (-1)^{|\theta|} f\big(a+\alpha\,(x-a)+\theta(y-x)\big)$$

in M as $j \to \infty$, and so, by virtue of (3.3) and the remarks preceding Theorem 2,

$$\begin{aligned} \mathrm{md}_{|\alpha|}(f_\alpha^a, I_x^y\lfloor\alpha) &= \Big\| \sum_{0\le\theta\le\alpha} (-1)^{|\theta|} f\big(a+\alpha\,(x-a)+\theta(y-x)\big)\Big\| \\ &\le \liminf_{j\to\infty} \Big\| \sum_{0\le\theta\le\alpha} (-1)^{|\theta|} f_j\big(a+\alpha\,(x-a)+\theta(y-x)\big)\Big\| \\ &= \liminf_{j\to\infty}\ \mathrm{md}_{|\alpha|}((f_j)_\alpha^a, I_x^y\lfloor\alpha). \end{aligned} \tag{4.14}$$

Arguing as in Step 2 of the proof of Theorem E, making use of the inequality (4.14), which coincides with (8.2) (see p. 44), and taking into account (2.4), we get:

$$\mathrm{TV}(f, I_a^b) \le \liminf_{j\to\infty}\ \mathrm{TV}(f_j, I_a^b) \le C.$$

This completes the proof of Theorem 2. □

Remark 4.2. Note that instead of condition (2.5) in Theorem 2 we may assume only that the value $c(a) = \sup_{j\in\mathbb{N}} \|f_j(a)\|$ is finite. In fact, it follows from Theorem B and condition (2.4) that, given $x \in I_a^b$ and $j \in \mathbb{N}$,

$$\|f_j(x)\| \le \|f_j(a)\| + \|f_j(x) - f_j(a)\| \le c(a) + \mathrm{TV}(f_j, I_a^x) \le c(a) + C.$$

5. Proof of Theorem B

In order to prove Theorem B, we need three more Lemmas 4, 5 and 6. The following equality will be needed in Lemma 4 (cf. [17, Part I, equality (3.4)]): given two multiindices $0 \le \beta \le \gamma \le 1$, we have:

$$|\{\alpha : \beta \le \alpha \le \gamma \text{ and } |\alpha| = i\}| = C_{|\gamma|-|\beta|}^{i-|\beta|} \quad \text{for all} \quad |\beta| \le i \le |\gamma|, \tag{5.1}$$

where $|A|$ denotes the number of elements in the set A and, given $0 \le j \le m$, $C_m^j = \binom{m}{j} = \frac{m!}{j!(m-j)!}$ is the usual binomial coefficient (with $0! = 1$). Also, recall (cf. Section 2) that a multiindex α is said to be even (odd) if $\alpha \in \mathcal{E}(n)$ ($\alpha \in \mathcal{O}(n)$, respectively).

Lemma 4. (a) *Given $m \in \mathbb{N}$ and integer $0 \le k \le m-1$, we have*:

$$\sum_{i \ge k/2}^{\le m/2} C_{m-k}^{2i-k} = \sum_{i \ge (k+1)/2}^{\le (m+1)/2} C_{m-k}^{2i-1-k} = 2^{m-k-1},$$

where the summations are taken over integer *i in the ranges specified.*

(b) *Given two multiindices $0 \le \beta \le \gamma \le 1$ with $\beta \ne \gamma$, we have*:

$$|\{\text{even } \alpha : \beta \le \alpha \le \gamma\}| = |\{\text{odd } \alpha : \beta \le \alpha \le \gamma\}|.$$

Proof. (a) Since the case $m = 1$ is clear, we suppose that $m \ge 2$. By the binomial formula, $0 = (1-1)^{m-k} = \sum_{j=0}^{m-k} (-1)^j C_{m-k}^j$, which is equal to

$$\sum_{j=0}^{(m-k-1)/2} C_{m-k}^{2j} - \sum_{j=1}^{(m-k+1)/2} C_{m-k}^{2j-1} \quad \text{if } m \text{ and } k \text{ are of different evenness,}$$

and

$$\sum_{j=0}^{(m-k)/2} C_{m-k}^{2j} - \sum_{j=1}^{(m-k)/2} C_{m-k}^{2j-1} \quad \text{if } m \text{ and } k \text{ are of the same evenness.}$$

If k is even, we change the summation index $j \mapsto i = (2j+k)/2$ in these sums and get, respectively:

$$0 = \sum_{i=k/2}^{(m-1)/2} C_{m-k}^{2i-k} - \sum_{i=(k+2)/2}^{(m+1)/2} C_{m-k}^{2i-1-k} \qquad \text{if } m \text{ is odd,}$$

$$0 = \sum_{i=k/2}^{m/2} C_{m-k}^{2i-k} - \sum_{i=(k+2)/2}^{m/2} C_{m-k}^{2i-1-k} \qquad \text{if } m \text{ is even,}$$

while if k is odd, we change the summation index $j \mapsto i = (2j+1+k)/2$ in the first sum and $j \mapsto i = (2j-1+k)/2$ in the second sum and get, respectively:

$$0 = \sum_{i=(k+1)/2}^{m/2} C_{m-k}^{2i-1-k} - \sum_{i=(k+1)/2}^{m/2} C_{m-k}^{2i-k} \qquad \text{if } m \text{ is even,}$$

$$0=\sum_{i=(k+1)/2}^{(m+1)/2} C_{m-k}^{2i-1-k}-\sum_{i=(k+1)/2}^{(m-1)/2} C_{m-k}^{2i-k} \qquad \text{if } m \text{ is odd.}$$

The second equality in (a) follows from the equality $(1+1)^{m-k}=\sum_{j=0}^{m-k} C_{m-k}^{j}$.

Remark. The first equality in Lemma 4(a) can be written also as

$$\sum_{i=[(k+2)/2]}^{[(m+1)/2]} C_{m-k}^{2i-1-k}=\sum_{i=[(k+1)/2]}^{[m/2]} C_{m-k}^{2i-k}, \quad m\in\mathbb{N}, \quad 0\le k\le m-1,$$

where $[r]$ designates the integer part of $r\in\mathbb{R}$. However, we prefer the equality in Lemma 4(a) since it is more simple and suggestive.

(b) By virtue of (5.1), the left-hand side of the equality is equal to

$$\bigl|\{\alpha:\beta\le\alpha\le\gamma \text{ and } |\alpha|=2i \text{ for all } i \text{ s.t. } |\beta|\le 2i\le|\gamma|\}\bigr|=\sum_{i\ge|\beta|/2}^{\le|\gamma|/2} C_{|\gamma|-|\beta|}^{2i-|\beta|},$$

and the right-hand side of the equality is equal to

$$\begin{aligned}&\bigl|\{\alpha:\beta\le\alpha\le\gamma \text{ and } |\alpha|=2i-1 \text{ for all } i \text{ s.t. } |\beta|\le 2i-1\le|\gamma|\}\bigr|\\ &\qquad=\sum_{i\ge(|\beta|+1)/2}^{\le(|\gamma|+1)/2} C_{|\gamma|-|\beta|}^{2i-1-|\beta|}.\end{aligned}$$

It remains to put $m=|\gamma|$ and $k=|\beta|$, note that $k<m$ and apply the equality from the previous assertion (a). □

If $(M,d,+)$ is a metric semigroup, then, by virtue of the triangle inequality for d and the translation invariance of metric d on M, we have, for all $u,v,u',v'\in M$:

$$\begin{aligned} d(u,v) &\le d(u',v')+d(u+u',v+v'),\\ d(u+u',v+v') &\le d(u,v)+d(u',v'). \end{aligned} \tag{5.2}$$

Inequality (5.2) yields that the addition operation $(u,v)\mapsto u+v$ is a continuous map from $M\times M$ into M. More generally, if $u_j\to u$, $v_j\to v$, $u_j'\to u'$ and $v_j'\to v'$ as $j\to\infty$ (convergence of sequences in M), then $\lim_{j\to\infty} d(u_j+v_j,u_j'+v_j')=d(u+v,u'+v')$.

Lemma 5. *If $m \in \mathbb{N}$, $u, v \in M$, $\{u_j\}_{j=1}^m$, $\{v_j\}_{j=1}^m \subset M$ and*

$$\sum_{i=1}^{\leq m/2} u_{2i} \;+\; u \;+\; \sum_{i=1}^{\leq (m+1)/2} v_{2i-1} = \sum_{i=1}^{\leq m/2} v_{2i} \;+\; v \;+\; \sum_{i=1}^{\leq (m+1)/2} u_{2i-1}, \tag{5.3}$$

then

$$d(u,v) \leq \sum_{j=1}^{m} d(u_j, v_j). \tag{5.4}$$

Proof. Observe that if $u + \ell_1 + \cdots + \ell_k = v + r_1 + \cdots + r_k$ for some $k \in \mathbb{N}$ and $\{\ell_i, r_i\}_{i=1}^k \subset M$, then $d(u,v) \leq \sum_{i=1}^k d(\ell_i, r_i)$. In fact, by the translation invariance of d and inequality (5.2), we have:

$$\begin{aligned} d(u,v) &= d\Big(u + \sum_{i=1}^k \ell_i, v + \sum_{i=1}^k \ell_i\Big) = d\Big(v + \sum_{i=1}^k r_i, v + \sum_{i=1}^k \ell_i\Big) \\ &= d\Big(\sum_{i=1}^k r_i, \sum_{i=1}^k \ell_i\Big) \leq \sum_{i=1}^k d(r_i, \ell_i). \end{aligned}$$

Applying this observation and equality (5.3), we get:

$$d(u,v) \leq \sum_{i=1}^{\leq m/2} d(u_{2i}, v_{2i}) + \sum_{i=1}^{\leq (m+1)/2} d(v_{2i-1}, u_{2i-1}) = \sum_{j=1}^{m} d(u_j, v_j). \qquad \square$$

Remark 5.1. In particular, (in)equalities (5.3) and (5.4) hold for *odd* m if

$$u + \sum_{i=1}^{(m-1)/2} u_{2i} = \sum_{i=1}^{(m+1)/2} u_{2i-1} \quad \text{and} \quad v + \sum_{i=1}^{(m-1)/2} v_{2i} = \sum_{i=1}^{(m+1)/2} v_{2i-1}, \tag{5.5}$$

and for *even* m if either

$$u + \sum_{i=1}^{m/2} u_{2i} = v + \sum_{i=1}^{m/2} u_{2i-1} \quad \text{and} \quad \sum_{i=1}^{m/2} v_{2i} = \sum_{i=1}^{m/2} v_{2i-1}, \tag{5.6}$$

or

$$\sum_{i=1}^{m/2} u_{2i} = v + \sum_{i=1}^{m/2} u_{2i-1} \quad \text{and} \quad \sum_{i=1}^{m/2} v_{2i} = u + \sum_{i=1}^{m/2} v_{2i-1}. \tag{5.7}$$

In the next lemma we set $\mathcal{A}_0 \equiv \mathcal{A}_0(n) = \{\theta \in \mathbb{N}_0^n : \theta \le 1\}$. Also, we stick to the following conventions: '$u \doteq 0$' will mean that u is omitted in the formula under consideration (especially in a metric semigroup with no zero), and a sum over the empty set is also omitted in any context (i.e., $\sum_{\varnothing} \doteq 0$).

Lemma 6. *Given a map $h : \mathcal{A}_0 \to M$ and a multiindex $\gamma \in \mathcal{A}_0$, we have*:

$$\sum_{\mathrm{ev}\,\alpha\le\gamma}\;\sum_{\mathrm{ev}\,\theta\le\alpha} h(\theta) = c_\gamma + \sum_{\mathrm{od}\,\alpha\le\gamma}\;\sum_{\mathrm{ev}\,\theta\le\alpha} h(\theta), \tag{5.8}$$

where $c_\gamma \doteq 0$ if γ is odd, and $c_\gamma = h(\gamma)$ if γ is even, and

$$\sum_{\mathrm{od}\,\alpha\le\gamma}\;\sum_{\mathrm{od}\,\theta\le\alpha} h(\theta) = d_\gamma + \sum_{\mathrm{ev}\,\alpha\le\gamma}\;\sum_{\mathrm{od}\,\theta\le\alpha} h(\theta), \tag{5.9}$$

where $d_\gamma = h(\gamma)$ if γ is odd, and $d_\gamma \doteq 0$ if γ is even.

Proof. 0. Denote by $\mathcal{L}$ (by $\mathcal{R}$) the set of all 'admissible' θ's at the left (right) hand side of the equality under consideration and, given $\theta \in \mathcal{L}$ (and $\theta \in \mathcal{R}$), by $L(\theta)$ (and by $R(\theta)$)—the multiplicity of the term $h(\theta)$ at the left (and right) hand sum(s). Then the equality can be rewritten as

$$\sum_{\theta\in\mathcal{L}} L(\theta)h(\theta) = \sum_{\theta\in\mathcal{R}} R(\theta)h(\theta), \tag{5.10}$$

where $L(\theta)h(\theta)$ denotes the sum of terms of the form $h(\theta)$ taken $L(\theta)$ times (and likewise for $R(\theta)h(\theta)$). In what follows in order to prove (5.10), we show that $\mathcal{L} = \mathcal{R}$ and $L(\theta) = R(\theta)$ for all $\theta \in \mathcal{L} = \mathcal{R}$.

We divide the proof into four steps for clarity.

In the first two steps we let γ be odd (i.e., $0 \le \gamma \le 1$ and $|\gamma|$ is odd).

1. Let us establish (5.8). We have $\mathcal{L} = \{\text{even } \theta : \exists\,\text{even } \alpha \le \gamma \text{ s.t. } \theta \le \alpha\}$, i.e., $\mathcal{L} = \{\text{even } \theta : \theta \le \gamma\}$, and $\mathcal{R} = \{\text{even } \theta : \exists\,\text{odd } \alpha \le \gamma\,\text{s.t.}\,\theta \le \alpha\}$. The sets $\mathcal{L}$ and $\mathcal{R}$ are nonempty ($0 \in \mathcal{L}$ and $0 \in \mathcal{R}$) and $\mathcal{L} = \mathcal{R}$. In fact, the inclusion $\mathcal{L} \supset \mathcal{R}$ is clear, and so, we let $\theta \in \mathcal{L}$. Since θ is even, γ is odd and $\theta \le \gamma$, there exists $i \in \{1, \dots, n\}$ s.t. $\theta_i = 0$ and $\gamma_i = 1$. We set $\alpha = (\theta_1, \dots, \theta_{i-1}, 1, \theta_{i+1}, \dots, \theta_n)$. It follows that $\alpha \le \gamma$, $|\alpha| = |\theta| + 1$ is odd and $\theta \le \alpha$, and so, $\theta \in \mathcal{R}$.

Given $\theta \in \mathcal{L} = \mathcal{R}$, we find $\theta \ne \gamma$, $L(\theta) = |\{\text{even } \alpha : \theta \le \alpha \le \gamma\}|$ and $R(\theta) = |\{\text{odd } \alpha : \theta \le \alpha \le \gamma\}|$. By Lemma 4(b), $L(\theta) = R(\theta)$, and so, (5.10) holds implying (5.8) with $c_\gamma \doteq 0$.

2. Let us prove (5.9). If $|\gamma| = 1$, then the equality is immediate: the left-hand side is equal to $h(\gamma) = d_\gamma$, while the double sum at the right is omitted (in fact, even $\alpha \le \gamma$ implies $\alpha = 0$, and so, no odd θ s.t. $\theta \le 0$ exists). Now, if $|\gamma| > 1$, then $\mathcal{L} = \{\text{odd } \theta : \theta \le \gamma\}$ and $\mathcal{R} = \{\text{odd } \theta : \exists\, \text{even } \alpha \le \gamma \,\text{s.t.}\, \theta \le \alpha\} \cup \{\gamma\}$ (disjoint union), and $\mathcal{L} = \mathcal{R}$. Let $\theta \in \mathcal{L} = \mathcal{R}$. If $\theta \ne \gamma$, then $L(\theta) = |\{\text{odd } \alpha : \theta \le \alpha \le \gamma\}|$ and $R(\theta) = |\{\text{even } \alpha : \theta \le \alpha \le \gamma\}|$, and so, by Lemma 4(b), $L(\theta) = R(\theta)$. Now if $\theta = \gamma$, then $L(\gamma) = |\{\text{odd } \alpha : \gamma \le \alpha \le \gamma\}| = 1$, and since $d_\gamma = h(\gamma)$, then $R(\gamma) = 1$ as well. The conclusion follows as in Step 1.

Suppose that γ is even.

3. In order to prove (5.8), we first note that if $\gamma = 0$, then the double sum at the right is omitted and the double sum at the left is equal to $h(0) = c_0$. Assume that $\gamma \ne 0$. Then $\mathcal{L} = \{\text{even } \theta : \theta \le \gamma\}$ and $\mathcal{R} = \{\gamma\} \cup \{\text{even } \theta : \exists\, \text{odd } \alpha \le \gamma \,\text{s.t.}\, \theta \le \alpha\}$ (disjoint union), and $\mathcal{L} = \mathcal{R}$. Let $\theta \in \mathcal{L} = \mathcal{R}$. Then $L(\theta) = |\{\text{even } \alpha : \theta \le \alpha \le \gamma\}|$ and, in particular, $L(\gamma) = 1$. If $\theta = \gamma$, then, since $c_\gamma = h(\gamma)$, we have $R(\gamma) = 1$, and if $\theta \ne \gamma$, then $R(\theta) = |\{\text{odd } \alpha : \theta \le \alpha \le \gamma\}|$, and so, by Lemma 4(b), $L(\theta) = R(\theta)$.

4. Finally, we prove (5.9). Since the equality is clear for $\gamma = 0$ (i.e., 'empty' equality), we assume that $|\gamma| > 0$. We have $\mathcal{L} = \{\text{odd } \theta : \theta \le \gamma\}$, $\mathcal{R} = \{\text{odd } \theta : \exists\, \text{even } \alpha \le \gamma \,\text{s.t.}\, \theta \le \alpha\}$ and $\mathcal{L} = \mathcal{R}$. Given $\theta \in \mathcal{L} = \mathcal{R}$, we find $\theta \ne \gamma$, $L(\theta) = |\{\text{odd } \alpha : \theta \le \alpha \le \gamma\}|$ and $R(\theta) = |\{\text{even } \alpha : \theta \le \alpha \le \gamma\}|$, and so, by Lemma 4(b), $L(\theta) = R(\theta)$. $\square$

Now we are in a position to prove Theorem B.

Proof of Theorem B. It suffices to prove only the first inequality: the second one follows from the first inequality, (2.2) and (2.3). Setting $u = f(x)$ and $v = f(y)$ and taking into account (3.4), the first inequality in Theorem B can be rewritten equivalently as

$$d(u, v) \le \sum_{0 \ne \alpha \le 1} d(u(\alpha), v(\alpha)) = \sum_{j=1}^{n} \sum_{|\alpha|=j} d(u(\alpha), v(\alpha)) \tag{5.11}$$

(the sum over $|\alpha| = j$ designates the sum over $0 \ne \alpha \le 1$ s.t. $|\alpha| = j$), where, given $\alpha, \theta \in \mathcal{A}_0$, we set $h(\theta) = f(x + \theta(y - x))$,

$$u(\alpha) = \sum_{\text{ev}\, \theta \le \alpha} h(\theta), \quad \text{and} \quad v(\alpha) = \sum_{\text{od}\, \theta \le \alpha} h(\theta) \quad \text{if } \alpha \ne 0 \quad \text{and} \quad v(0) \doteq 0.$$

In order to establish (5.11), given integer $0 \le j \le n$, we also set

$$u_j = \sum_{|\alpha|=j} u(\alpha) \qquad \text{and} \qquad v_j = \sum_{|\alpha|=j} v(\alpha)$$

and note that

$$u_0 = u(0) = h(0) = u,\ v_0 = v(0) \doteq 0,\ v = h(1),\ u_n = u(1) \text{ and } v_n = v(1).$$

Suppose that we have already verified equalities (5.5) if $m = n$ is odd and (5.6) if $m = n$ is even. Applying Lemma 5, we get inequality (5.4), where, by virtue of (5.2), we have:

$$d(u_j, v_j) = d\biggl(\sum_{|\alpha|=j} u(\alpha), \sum_{|\alpha|=j} v(\alpha)\biggr) \le \sum_{|\alpha|=j} d(u(\alpha), v(\alpha)).$$

Now, (5.11) follows if we sum these inequalities over $j = 1, \dots, n$ and take into account (5.4).

It remains to verify equalities (5.5) and (5.6). For this, we apply Lemma 6 with $\gamma = 1$ and note that $m = n = |\gamma| = |1|$. Suppose that $n = |1|$ is odd. By virtue of (5.8), we have:

$$\begin{aligned} u + \sum_{i=1}^{(m-1)/2} u_{2i} &= \sum_{i=0}^{(n-1)/2} u_{2i} = \sum_{i=0}^{(n-1)/2} \sum_{|\alpha|=2i} u(\alpha) = \sum_{\mathrm{ev}\,\alpha\le 1} u(\alpha) \\ &= \sum_{\mathrm{od}\,\alpha\le 1} u(\alpha) = \sum_{i=1}^{(n+1)/2} \sum_{|\alpha|=2i-1} u(\alpha) = \sum_{i=1}^{(m+1)/2} u_{2i-1}, \end{aligned}$$

and by virtue of (5.9), we get:

$$\begin{aligned} v + \sum_{i=1}^{(m-1)/2} v_{2i} &= h(1) + \sum_{i=0}^{(n-1)/2} v_{2i} = h(1) + \sum_{i=0}^{(n-1)/2} \sum_{|\alpha|=2i} v(\alpha) = h(1) + \sum_{\mathrm{ev}\,\alpha\le 1} v(\alpha) \\ &= \sum_{\mathrm{od}\,\alpha\le 1} v(\alpha) = \sum_{i=1}^{(n+1)/2} \sum_{|\alpha|=2i-1} v(\alpha) = \sum_{i=1}^{(m+1)/2} v_{2i-1}, \end{aligned}$$

which establishes (5.5). Now suppose that $n = |1|$ is even. By (5.8), we get:

$$\begin{aligned} u + \sum_{i=1}^{m/2} u_{2i} &= \sum_{i=0}^{n/2} u_{2i} = \sum_{i=0}^{n/2} \sum_{|\alpha|=2i} u(\alpha) = \sum_{\mathrm{ev}\,\alpha\le 1} u(\alpha) \\ &= h(1) + \sum_{\mathrm{od}\,\alpha\le 1} u(\alpha) = v + \sum_{i=1}^{n/2} \sum_{|\alpha|=2i-1} u(\alpha) = v + \sum_{i=1}^{m/2} u_{2i-1}, \end{aligned}$$

and by virtue of (5.9), we have:

$$\begin{aligned} \sum_{i=1}^{m/2} v_{2i} &= \sum_{i=0}^{n/2} v_{2i} = \sum_{i=0}^{n/2} \sum_{|\alpha|=2i} v(\alpha) = \sum_{\mathrm{ev}\,\alpha\le 1} v(\alpha) \\ &= \sum_{\mathrm{od}\,\alpha\le 1} v(\alpha) = \sum_{i=1}^{n/2} \sum_{|\alpha|=2i-1} v(\alpha) = \sum_{i=1}^{m/2} v_{2i-1}, \end{aligned}$$

which establishes (5.6) and completes the proof of Theorem B. $\square$

Remark 5.2. The left-hand side inequality in Theorem B is of interest when $x < y$. However, if $x \le y$ and $x \not< y$, it can be refined in the following way (cf. [17, Part I, Lemma 6]): given $x, y \in I_a^b$, $x < y$, and $0 \ne \gamma \le 1$, we have:

$$d\big(f(x), f(x+\gamma(y-x))\big) \le \sum_{0\ne\alpha\le\gamma} \mathrm{md}_{|\alpha|}(f_\alpha^x, I_x^y\lfloor\alpha).$$

In fact, by Theorem B, we find

$$d\big(f(x), f(x+\gamma(y-x))\big) \le \sum_{0\ne\alpha\le 1} \mathrm{md}_{|\alpha|}(f_\alpha^x, I_x^{x+\gamma(y-x)}\lfloor\alpha),$$

where, by virtue of (3.4), the mixed difference at the right is equal to

$$d\bigg(\sum_{\mathrm{ev}\,\theta\le\alpha} f(x+\theta\gamma(y-x)), \sum_{\mathrm{od}\,\overline{\theta}\le\alpha} f(x+\overline{\theta}\gamma(y-x))\bigg). \tag{5.12}$$

If $\alpha \not\le \gamma$, then $\alpha_i = 1$ and $\gamma_i = 0$ for some $i \in \{1, \dots, n\}$, and so, arguing as in Remark 2.1 we find $x + \theta\gamma(y-x) = x + \overline{\theta}\gamma(y-x)$ for all even θ with $\theta \le \alpha$ implying that (5.12) is equal to zero. Now if $\alpha \le \gamma$, then $\theta\gamma = \theta$ for any $\theta \le \alpha$, and so, (5.12) coincides with the right-hand side of (3.4).

6. Proof of Lemma 2

In order to prove Lemma 2, we need an auxiliary Lemma 7, which plays the same role as Lemma 6 above.

Lemma 7. *Given a map $h : \mathcal{A}_0 \to M$ and a multiindex $\alpha \in \mathcal{A}_0$, we have: if $1-\alpha$ is even, then the following two equalities hold:*

$$\sum_{\mathrm{ev}\,\theta\le\alpha} h(1-\alpha+\theta) + \sum_{\mathrm{od}\,\beta\le 1-\alpha}\ \sum_{\mathrm{ev}\,\theta\le\alpha+\beta} h(\theta) = \sum_{\mathrm{ev}\,\beta\le 1-\alpha}\ \sum_{\mathrm{ev}\,\theta\le\alpha+\beta} h(\theta), \tag{6.1}$$

$$\sum_{\mathrm{od}\,\theta\le\alpha} h(1-\alpha+\theta) + \sum_{\mathrm{od}\,\beta\le 1-\alpha}\ \sum_{\mathrm{od}\,\theta\le\alpha+\beta} h(\theta) = \sum_{\mathrm{ev}\,\beta\le 1-\alpha}\ \sum_{\mathrm{od}\,\theta\le\alpha+\beta} h(\theta), \tag{6.2}$$

and if $1-\alpha$ is odd, then the following two equalities hold:

$$\sum_{1-\alpha\le\mathrm{ev}\,\theta\le 1} h(\theta) + \sum_{\mathrm{ev}\,\beta\le 1-\alpha}\ \sum_{\mathrm{ev}\,\theta\le\alpha+\beta} h(\theta) = \sum_{\mathrm{od}\,\beta\le 1-\alpha}\ \sum_{\mathrm{ev}\,\theta\le\alpha+\beta} h(\theta), \tag{6.3}$$

$$\sum_{1-\alpha\le\mathrm{od}\,\theta\le 1} h(\theta) + \sum_{\mathrm{ev}\,\beta\le 1-\alpha}\ \sum_{\mathrm{od}\,\theta\le\alpha+\beta} h(\theta) = \sum_{\mathrm{od}\,\beta\le 1-\alpha}\ \sum_{\mathrm{od}\,\theta\le\alpha+\beta} h(\theta). \tag{6.4}$$

Proof. As in the proof of Lemma 6, the main idea is to establish equality (5.10). We divide the proof into four steps.

Suppose that $1-\alpha$ is even.

1. Let us prove (6.1). If $\alpha = 1$, then $1-\alpha = 0$ is even, and equality (6.1) is equivalent to the identity $\sum_{\mathrm{ev}\,\theta\le 1} h(\theta) + 0 = \sum_{\mathrm{ev}\,\theta\le 1} h(\theta)$. If $\alpha = 0$ and if $1-\alpha = 1$ is even, then (6.1) can be written as

$$h(1) + \sum_{\mathrm{od}\,\beta\le 1}\ \sum_{\mathrm{ev}\,\theta\le\beta} h(\theta) = \sum_{\mathrm{ev}\,\beta\le 1}\ \sum_{\mathrm{ev}\,\theta\le\beta} h(\theta),$$

which was established in (5.8) for even $\gamma = 1$. Thus, in what follows we assume that $\alpha \ne 0, 1$, i.e., $0 < |\alpha| < n$.

We have $\mathcal{L} = \mathcal{L}_1 \cup \mathcal{L}_2$, where $\mathcal{L}_1 = \{1-\alpha+\theta' : \exists\,\text{even } \theta' \le \alpha\}$ (so that $1-\alpha \in \mathcal{L}_1$) and $\mathcal{L}_2 = \{\text{even } \theta : \exists\,\text{odd } \beta \le 1-\alpha \text{ s.t. } \theta \le \alpha+\beta\}$ (so that $0 \in \mathcal{L}_2$), and $\mathcal{R} = \{\text{even } \theta : \exists\,\text{even } \beta \le 1-\alpha \text{ s.t. } \theta \le \alpha+\beta\}$, i.e., $\mathcal{R} = \{\text{even } \theta : \theta \le 1\}$. We are going to show that $\mathcal{L} = \mathcal{R}$. This equality follows immediately from the definition of $\mathcal{R}$ and the following two assertions:

$$\theta \in \mathcal{L}_1 \iff \theta \text{ is even and } \alpha \vee \theta = 1, \tag{6.5}$$

$$\theta \in \mathcal{L}_2 \iff \theta \text{ is even and } \alpha \vee \theta \ne 1, \tag{6.6}$$

where $\alpha \vee \theta \equiv \max\{\alpha, \theta\} = \alpha + \theta - \alpha\,\theta$; in particular, (6.5) and (6.6) imply that $\mathcal{L}_1$ and $\mathcal{L}_2$ are disjoint. Let us prove (6.5). If $\theta \in \mathcal{L}_1$, then $\theta = 1 - \alpha + \theta'$ for some even $\theta' \leq \alpha$ and, since $1 - \alpha$ is even and $|\theta| = |1 - \alpha| + |\theta'|$, then θ is even, $\theta \leq (1 - \alpha) + \alpha = 1$ and

$$\alpha \vee \theta = \alpha + (1 - \alpha + \theta') - \alpha(1 - \alpha + \theta') = 1 + \theta' - \alpha\,\theta' = 1.$$

Conversely, if θ is even and $\alpha \vee \theta = 1$, then $\alpha + \theta - \alpha\theta = 1$ or $\alpha + \theta = 1 + \alpha\theta \geq 1$. Setting $\theta' = \alpha + \theta - 1$, we find $\theta = 1 - \alpha + \theta'$, where $|\theta'| = |\alpha| + |\theta| - n = |\theta| - |1 - \alpha|$ is even and $\theta' \leq \alpha$, and so, $\theta \in \mathcal{L}_1$. Now we establish (6.6). If $\theta \in \mathcal{L}_2$, then θ is even and there exists odd $\beta \leq 1 - \alpha$ s.t. $\theta \leq \alpha + \beta$, and so, $\alpha \leq \alpha + \beta$ and $\theta \leq \alpha + \beta$ imply $\alpha \vee \theta \leq \alpha + \beta$. Since β is odd, $1 - \alpha$ is even and $\beta \leq 1 - \alpha$, we have $|\beta| < |1 - \alpha| = n - |\alpha|$. It follows that

$$|\alpha \vee \theta| \leq |\alpha + \beta| = |\alpha| + |\beta| < |\alpha| + (n - |\alpha|) = n,$$

and so, $\alpha \vee \theta \neq 1$. Conversely, if θ is even and $\alpha \vee \theta \neq 1$, then there exists $i \in \{1, \ldots, n\}$ s.t. $\alpha_i = 0$ and $\theta_i = 0$. Setting $\beta = (\beta_1, \ldots, \beta_n)$ with $\beta_i = 0$ and $\beta_j = 1 - \alpha_j$ if $j \neq i$, we find $\beta \leq 1 - \alpha$, $|\beta| = |1 - \alpha| - 1$ is odd and $\theta \leq \alpha + \beta$, and so, $\theta \in \mathcal{L}_2$.

In order to calculate the values $L(\theta)$ and $R(\theta)$ for $\theta \in \mathcal{L} = \mathcal{R}$, we note that, given $0 \leq \beta \leq 1 - \alpha$, we have:

$$\theta \leq \alpha + \beta \quad \text{is equivalent to} \quad (1 - \alpha)\theta \leq \beta. \tag{6.7}$$

In fact, condition $0 \leq \beta \leq 1 - \alpha$ is equivalent to condition $\alpha\beta = 0$:

$$0 \leq \beta \leq 1 - \alpha \iff \beta(1 - \alpha) = \beta \iff \beta - \alpha\beta = \beta \iff \alpha\beta = 0,$$

and so, if $\theta \leq \alpha + \beta$, then $(1 - \alpha)\theta \leq (1 - \alpha)(\alpha + \beta) = (1 - \alpha)\alpha + \beta - \alpha\beta = \beta$, and if $(1 - \alpha)\theta \leq \beta$, then $\theta - \alpha\theta \leq \beta$, and so, $\theta \leq \alpha\theta + \beta \leq \alpha + \beta$.

Given $\theta \in \mathcal{R}$, by virtue of (6.7), we find

$$R(\theta) = |\{\text{even } \beta : \beta \leq 1 - \alpha \text{ and } \theta \leq \alpha + \beta\}| = |\{\text{even } \beta : (1 - \alpha)\theta \leq \beta \leq 1 - \alpha\}|.$$

If $\theta \in \mathcal{L}_1$, then there exists a unique even $\theta' \leq \alpha$ s.t. $\theta = 1 - \alpha + \theta'$, and so, since $\theta \notin \mathcal{L}_2$, then $L(\theta) = 1$. At the same time,

$$(1 - \alpha)\theta = (1 - \alpha)(1 - \alpha + \theta') = (1 - \alpha)^2 + (1 - \alpha)\theta' = 1 - \alpha,$$

and so, by the above, $R(\theta) = 1$ as well. Suppose now that $\theta \in \mathcal{L}_2$. Then, by (6.6), $1 \neq \alpha \vee \theta = \alpha + \theta - \alpha\theta = \alpha + (1-\alpha)\theta$ or $(1-\alpha)\theta \neq 1-\alpha$, and so, taking into account (6.7) and Lemma 4(b) we find that

$$L(\theta) = |\{\text{odd } \beta : \beta \leq 1-\alpha \text{ and } \theta \leq \alpha+\beta\}| = |\{\text{odd } \beta : (1-\alpha)\theta \leq \beta \leq 1-\alpha\}|$$

is equal to $R(\theta)$.

In the rest of the proof we exhibit only the essential ingredients and differences.

2. Let us establish (6.2). If $\alpha = 1$, we get an identity, and if $\alpha = 0$ and $1 = 1-\alpha$ is even, we get equality (5.9) with even $\gamma = 1$, and so, we suppose that $0 < |\alpha| < n$. We have $\mathcal{L} = \mathcal{L}_1 \cup \mathcal{L}_2$, where $\mathcal{L}_1 = \{1-\alpha+\theta' : \exists \text{ odd } \theta' \leq \alpha\}$ and $\mathcal{L}_2 = \{\text{odd } \theta : \exists \text{ odd } \beta \leq 1-\alpha \text{ s.t. } \theta \leq \alpha+\beta\}$, and $\mathcal{R} = \{\text{odd } \theta : \exists \text{ even } \beta \leq 1-\alpha \text{ s.t. } \theta \leq \alpha+\beta\}$, which, actually, is $\mathcal{R} = \{\text{odd } \theta : \theta \leq 1\}$. We need to verify only that $\mathcal{L}_1$ and $\mathcal{L}_2$ are nonempty: the rest of the proof of (6.2) (including (6.5) and (6.6)) is the same as in Step 1 where 'even θ' is replaced by 'odd θ'.

Since $\alpha \neq 0$, there exists $i \in \{1,\dots,n\}$ s.t. $\alpha_i = 1$, and so, if we set $\theta' = (\theta'_1,\dots,\theta'_n)$ with $\theta'_i = 1$ and $\theta'_j = 0$ if $j \neq i$, then $|\theta'| = 1$ is odd and $\theta' \leq \alpha$. It follows that $1-\alpha+\theta' \in \mathcal{L}_1$.

Since $\alpha \neq 1$, there exists $i \in \{1,\dots,n\}$ s.t. $\alpha_i = 0$, and so if we set $\beta = (\beta_1,\dots,\beta_n)$ with $\beta_i = 0$ and $\beta_j = 1-\alpha_j$ if $j \neq i$, then $|\beta| = |1-\alpha|-1$ is odd and $\beta \leq 1-\alpha$. Given $k \in \{1,\dots,n\}$, $k \neq i$, setting $\theta = (\theta_1,\dots,\theta_n)$ with $\theta_k = 1$ and $\theta_j = 0$ if $j \neq k$, we find $|\theta| = 1$ is odd and $\theta \leq \alpha+\beta$, and so, $\theta \in \mathcal{L}_2$.

Assume now that $1-\alpha$ is odd. Note that $\alpha \neq 1$.

3. Let us prove (6.3). If $\alpha = 0$ and $1 = 1-\alpha$ is odd, then (since $\operatorname{ev}\theta{=}1$ cannot hold in the first sum at the left of (6.3)) equality (6.3) is equivalent to (5.8) with odd $\gamma = 1$. Thus, we assume that $|\alpha| > 0$.

We have $\mathcal{L} = \mathcal{L}_1 \cup \mathcal{L}_2$, where $\mathcal{L}_1 = \{\text{even } \theta : 1-\alpha \leq \theta \leq 1\}$ and $\mathcal{L}_2 = \{\text{even } \theta : \exists \text{ even } \beta \leq 1-\alpha \text{ s.t. } \theta \leq \alpha+\beta\}$, and $\mathcal{R} = \{\text{even } \theta : \exists \text{ odd } \beta \leq 1-\alpha \text{ s.t. } \theta \leq \alpha+\beta\}$, and so, $\mathcal{R} = \{\text{even } \theta : \theta \leq 1\}$. We have to show that $\mathcal{L} = \mathcal{R}$.

First, we show that $\mathcal{L}_1$ and $\mathcal{L}_2$ are nonempty. Since $\alpha \neq 0$, $\alpha_i = 1$ for some $i \in \{1,\dots,n\}$, and so, setting $\theta = (\theta_1,\dots,\theta_n)$ with $\theta_i = 1$ and $\theta_j = 1-\alpha_j$ if $j \neq i$, we find that $1-\alpha \leq \theta \leq 1$ and $|\theta| = |1-\alpha|+1$ is even, whence $\theta \in \mathcal{L}_1$. Now, since $\alpha \neq 1$, $\alpha_i = 0$ for some $i \in \{1,\dots,n\}$, and if we set

$\beta = (\beta_1, \dots, \beta_n)$ with $\beta_i = 0$ and $\beta_j = 1 - \alpha_j$ if $j \neq i$, then we find that $|\beta| = |1 - \alpha| - 1$ is even, $\theta = 0$ is even and $0 \leq \alpha + \beta$, and so, $0 \in \mathcal{L}_2$.

Second, we assert that (6.5) and (6.6) hold; this will imply that $\mathcal{L}_1$ and $\mathcal{L}_2$ are disjoint and $\mathcal{L} = \mathcal{R}$. In order to prove (6.5), we let $\theta \in \mathcal{L}_1$. Then θ is even and $1 - \alpha \leq \theta \leq 1$, and so,

$$\alpha \vee \theta = \alpha + \theta - \alpha\theta = \alpha + (1 - \alpha)\theta = \alpha + (1 - \alpha) = 1.$$

Conversely, if θ is even and $\alpha \vee \theta = 1$, then $\alpha + \theta - \alpha\theta = 1$, and so, $(1 - \alpha)\theta = 1 - \alpha$ implying $1 - \alpha \leq \theta$ and $\theta \in \mathcal{L}_1$. The proof of (6.6) follows the same lines as in Step 1 if 'odd β' is replaced by 'even β'.

Given $\theta \in \mathcal{R}$, taking into account (6.7), we have $R(\theta) = |\{\text{odd } \beta : (1 - \alpha)\theta \leq \beta \leq 1 - \alpha\}|$. If $\theta \in \mathcal{L}_1$, then $\theta \notin \mathcal{L}_2$, and so, $L(\theta) = 1$; in this case $1 - \alpha \leq \theta$, and so, $(1 - \alpha)\theta = 1 - \alpha$ and $R(\theta) = 1$. Now if $\theta \in \mathcal{L}_2$, then $\alpha \vee \theta \neq 1$, and so, $(1 - \alpha)\theta \neq 1 - \alpha$ and, by virtue of Lemma 4(b), the value $L(\theta) = |\{\text{even } \beta : (1 - \alpha)\theta \leq \beta \leq 1 - \alpha\}|$ is equal to $R(\theta)$.

4. Finally, we establish (6.4). If $\alpha = 0$ and $1 = 1 - \alpha$ is odd, we get equality (5.9) with odd $\gamma = 1$. Assume that $|\alpha| > 0$. We have $\mathcal{L} = \mathcal{L}_1 \cup \mathcal{L}_2$, where $\mathcal{L}_1 = \{\text{odd } \theta : 1 - \alpha \leq \theta \leq 1\}$ (and so, $1 - \alpha \in \mathcal{L}_1$) and $\mathcal{L}_2 = \{\text{odd } \theta : \exists\, \text{even } \beta \leq 1 - \alpha \text{ s.t. } \theta \leq \alpha + \beta\}$, and $\mathcal{R} = \{\text{odd } \theta : \exists\, \text{odd } \beta \leq 1 - \alpha \text{ s.t. } \theta \leq \alpha + \beta\}$, and so, $\mathcal{R} = \{\text{odd } \theta : \theta \leq 1\}$. That $\mathcal{L}_2$ is nonempty can be seen as follows. Since $\alpha \neq 1$, $\alpha_i = 0$ for some $i \in \{1, \dots, n\}$, and so, if we set $\beta = (\beta_1, \dots, \beta_n)$ with $\beta_i = 0$ and $\beta_j = 1 - \alpha_j$ if $j \neq i$, then $\beta \leq 1 - \alpha$ and $|\beta| = |1 - \alpha| - 1$ is even. Now, since $\alpha \neq 0$, $\alpha_k = 1$ for some $k \neq i$. If we set $\theta = (\theta_1, \dots, \theta_n)$ with $\theta_k = 1$ and $\theta_j = 0$ if $j \neq k$, then $|\theta| = 1$ is odd and $\theta \leq \alpha + \beta$, and so, $\theta \in \mathcal{L}_2$. Assertion (6.5) with 'θ is even' replaced by 'θ is odd' is established as in Step 3, while the proof of (6.6) follows the same lines as in Step 1 with 'odd β' replaced by 'even β'. It follows that $\mathcal{L} = \mathcal{R}$. The proof completes with the last paragraph of Step 3. $\square$

Proof of Lemma 2. The inequality (actually, equality) is clear if $\alpha = 1$, and so, we assume that $\alpha \neq 1$. The mixed difference at the left-hand side of the inequality is given by (3.4), while given $\alpha \leq \beta \leq 1$, noting that $\alpha\beta = \alpha$ and applying equality (3.3) we get the following expression for the mixed difference at the right-hand side (cf. [17, Part I, expression (3.7)]):

$$\mathrm{md}_{|\beta|}\big(f_\beta^a, I_{a+\alpha(x-a)}^{x+\alpha(y-x)} \lfloor \beta\big) = d\Bigg(\sum_{\mathrm{ev}\,\theta \leq \beta} h(\theta), \sum_{\mathrm{od}\,\theta \leq \beta} h(\theta)\Bigg),$$

where $h(\theta) = f\big(a + (\alpha \vee \theta)(x - a) + \alpha\theta(y - x)\big)$ and $\alpha \vee \theta = \alpha + \theta - \alpha\theta$. Changing the summation multiindex $\beta \mapsto \beta - \alpha$ in the sum at the right of the inequality in Lemma 2, we find that it is equivalent to

$$d(u, v) \le \sum_{0 \le \beta \le 1-\alpha} d\Bigg(\sum_{\mathrm{ev}\,\theta \le \alpha+\beta} h(\theta), \sum_{\mathrm{od}\,\theta \le \alpha+\beta} h(\theta) \Bigg),$$

where

$$u = \sum_{\mathrm{ev}\,\theta \le \alpha} f(x + \theta(y - x)) \qquad \text{and} \qquad v = \sum_{\mathrm{od}\,\theta \le \alpha} f(x + \theta(y - x)).$$

Setting

$$u(\beta) = \sum_{\mathrm{ev}\,\theta \le \alpha+\beta} h(\theta) \quad \text{and} \quad v(\beta) = \sum_{\mathrm{od}\,\theta \le \alpha+\beta} h(\theta) \quad \text{if} \quad 0 \le \beta \le 1 - \alpha,$$

the last inequality can be rewritten as

$$d(u, v) \le \sum_{0 \le \beta \le 1-\alpha} d(u(\beta), v(\beta)) = \sum_{j=0}^{|1-\alpha|} \sum_{|\beta|=j} d(u(\beta), v(\beta)). \tag{6.8}$$

In order to establish (6.8), we will apply Lemma 5 with $m = |1-\alpha| + 1 = n - |\alpha| + 1$ and

$$u_j = \sum_{|\beta|=j-1} u(\beta) \quad \text{and} \quad v_j = \sum_{|\beta|=j-1} v(\beta) \quad \text{if} \quad 1 \le j \le m.$$

Suppose that we have already verified equalities (5.5) and (5.7). Then by Lemma 5, we get inequality (5.4), where, by virtue of (5.2),

$$d(u_j, v_j) = d\Bigg(\sum_{|\beta|=j-1} u(\beta), \sum_{|\beta|=j-1} v(\beta) \Bigg) \le \sum_{|\beta|=j-1} d(u(\beta), v(\beta)), \quad 1 \le j \le m.$$

Summing over $j = 1, \dots, m$ and taking into account (5.4), we arrive at (6.8):

$$d(u, v) \le \sum_{j=1}^{m} d(u_j, v_j) \le \sum_{j=1}^{|1-\alpha|+1} \sum_{|\beta|=j-1} d(u(\beta), v(\beta)).$$

Assume that $1-\alpha$ is even; then m is odd. Let us verify the first equality in (5.5). For this, we apply equality (6.1) and calculate the first sum at the left-hand side of (6.1). Given even $\theta \leq \alpha$, we have $1-\alpha+\theta \in \mathcal{L}_1$ (cf. Step 1 in the proof of Lemma 7), and so, by (6.5), $\alpha \vee (1-\alpha+\theta) = 1$ and $\alpha(1-\alpha+\theta) = \theta$, so that the definition of $h(1-\alpha+\theta)$ implies

$$\sum_{\mathrm{ev}\,\theta\leq\alpha} h(1-\alpha+\theta) = \sum_{\mathrm{ev}\,\theta\leq\alpha} f(x+\theta(y-x)) = u.$$

Applying equality (6.1), we get:

$$\begin{aligned}
u + \sum_{i=1}^{(m-1)/2} u_{2i} &= u + \sum_{i=1}^{|1-\alpha|/2} \sum_{|\beta|=2i-1} u(\beta) = u + \sum_{\mathrm{od}\,\beta\leq 1-\alpha} u(\beta) \\
&= \sum_{\mathrm{ev}\,\beta\leq 1-\alpha} u(\beta) = \sum_{i=0}^{|1-\alpha|/2} \sum_{|\beta|=2i} u(\beta) = \sum_{i=0}^{|1-\alpha|/2} u_{2i+1} \\
&= \sum_{i=1}^{(|1-\alpha|+2)/2} u_{2i-1} = \sum_{i=1}^{(m+1)/2} u_{2i-1},
\end{aligned}$$

and the first equality in (5.5) follows. In a similar manner we find that the first sum at the left-hand side of (6.2) is equal to v, and, by virtue of (6.2), the calculations above show that the second equality in (5.5) holds as well.

Now suppose that $1-\alpha$ is odd, and so, m (defined above) is even. In order to verify the first equality in (5.7), we calculate the first sum at the left-hand side of (6.3). Given even θ with $1-\alpha \leq \theta \leq 1$, we have (cf. Step 3 in the proof of Lemma 7) $\theta \in \mathcal{L}_1$ and $\alpha \vee \theta = 1$. Moreover (cf. [17, Part I, assertion (3.9)]), there exists a unique $\theta' \in \mathcal{A}_0$ s.t. $\theta' \leq \alpha$ and $\theta = 1-\alpha+\theta'$ (define θ' by $\theta' = \alpha+\theta-1$). Since $|\theta'| = |\alpha|+|\theta|-n = |\theta|-|1-\alpha|$ and $1-\alpha$ is odd, then θ' is odd, and $\alpha\theta = \alpha(1-\alpha+\theta') = \theta'$. It follows that $h(\theta) = f(x+\theta'(y-x))$. Changing the summation multiindex $\theta \mapsto \theta'$ in the first sum at the left of (6.3), we get:

$$\sum_{1-\alpha\leq\mathrm{ev}\,\theta\leq 1} h(\theta) = \sum_{\mathrm{od}\,\theta'\leq\alpha} f(x+\theta'(y-x)) = v.$$

Applying equality (6.3), we find

$$
\begin{aligned}
\sum_{i=1}^{m/2} u_{2i} &= \sum_{i=1}^{(|1-\alpha|+1)/2} \sum_{|\beta|=2i-1} u(\beta) = \sum_{\mathrm{od}\,\beta\le 1-\alpha} u(\beta) \\
&= v + \sum_{\mathrm{ev}\,\beta\le 1-\alpha} u(\beta) = v + \sum_{i=0}^{(|1-\alpha|-1)/2} \sum_{|\beta|=2i} u(\beta) = v + \sum_{i=0}^{(|1-\alpha|-1)/2} u_{2i+1} \\
&= v + \sum_{i=1}^{(|1-\alpha|+1)/2} u_{2i-1} = v + \sum_{i=1}^{m/2} u_{2i-1},
\end{aligned}
$$

which proves the first equality in (5.7). Similarly, the first sum at the left-hand side of (6.4) is equal to u, and, by virtue of (6.4), the calculations above prove the second equality in (5.7).

This completes the proof of Lemma 2. □

7. Proof of Lemma 3

Note that if $\mathcal{P} = \{x[\sigma]\}_{\sigma=0}^{\kappa}$ is a net partition of I_a^b, then

$$
I_a^b = \bigcup_{1\le\sigma\le\kappa} I_{x[\sigma-1]}^{x[\sigma]} = \bigcup_{1\le\sigma\le\kappa} \prod_{i=1}^{n} [x_i(\sigma_i - 1), x_i(\sigma_i)] = \prod_{i=1}^{n} \Biggl(\bigcup_{l=1}^{\kappa_i} I_{x_i(l-1)}^{x_i(l)} \Biggr) \tag{7.1}
$$

is a union of non-overlapping non-degenerated rectangles $I_{x[\sigma-1]}^{x[\sigma]}$ with the sides parallel to the coordinate axes. In this section it will be convenient and brief to term the union as in (7.1) also a *partition* of I_a^b (by non-overlapping non-degenerated rectangles).

If $\mathcal{P} = \{x[\sigma]\}_{\sigma=0}^{\kappa}$ and $\mathcal{P}' = \{x'[\sigma']\}_{\sigma'=0}^{\kappa'}$ are two net partitions of I_a^b, we say that $\mathcal{P}'$ is a *refinement* of $\mathcal{P}$ if $\mathcal{P} \subset \mathcal{P}'$. Also, for the sake of convenience we define the *n-th prevariation* of $f : I_a^b \to M$, corresponding to $\mathcal{P}$, by

$$
\mathrm{v}_n(f;\mathcal{P}) = \sum_{1\le\sigma\le\kappa} \mathrm{md}_n(f, I_{x[\sigma-1]}^{x[\sigma]}).
$$

It follows that the Vitali-type n-th variation of f is given by $V_n(f, I_a^b) = \sup_{\mathcal{P}} \mathrm{v}_n(f;\mathcal{P})$, where the supremum is taken over all net partitions $\mathcal{P}$ of I_a^b.

The basic ingredient in the proof of Lemma 3 is the following

Lemma 8. *Given $f : I_a^b \to M$, if $\mathcal{P}$ and $\mathcal{P}'$ are two net partitions of I_a^b s.t. $\mathcal{P} \subset \mathcal{P}'$, then $\mathrm{v}_n(f;\mathcal{P}) \le \mathrm{v}_n(f;\mathcal{P}')$.*

In order to prove this lemma we need three more Lemmas 9–11. In what follows we fix a map $f : I_a^b \to M$.

Lemma 9. *Given $x, y \in I_a^b$ with $x < y$ and $x' \in I_a^b$, we have the following partition of I_x^y, induced by the point x':*

$$I_x^y = \bigcup_{1-\xi \le \alpha \le 1} I_{x+\alpha\xi(x'-x)}^{x'+\alpha(y-x')}, \tag{7.2}$$

where the multiindex $\xi \equiv \xi(x, x', y) = (\xi_1, \dots, \xi_n)$ is given by

$$\xi_i \equiv \xi_i(x, x', y) = \begin{cases} 1 & \text{if } \; x_i < x_i' < y_i, \\ 0 & \text{if } \; x_i' \le x_i \text{ or } x_i' \ge y_i, \end{cases} \qquad i \in \{1, \dots, n\}, \tag{7.3}$$

and

$$\mathrm{md}_n(f, I_x^y) \le \sum_{1-\xi \le \alpha \le 1} \mathrm{md}_n\big(f, I_{x+\alpha\xi(x'-x)}^{x'+\alpha(y-x')}\big). \tag{7.4}$$

Before we prove Lemma 9, let us establish two of its particular variants as Lemmas 10 and 11 (note that in Lemma 10 the rectangles in the union may degenerate).

Lemma 10. *If $x, y \in I_a^b$ with $x < y$ and $x' \in I_x^y$, then we have the following union of non-overlapping (possibly, degenerated) rectangles*

$$I_x^y = \bigcup_{0 \le \alpha \le 1} I_{x+\alpha(x'-x)}^{x'+\alpha(y-x')}, \tag{7.5}$$

and the following inequality holds:

$$\mathrm{md}_n(f, I_x^y) \le \sum_{0 \le \alpha \le 1} \mathrm{md}_n\big(f, I_{x+\alpha(x'-x)}^{x'+\alpha(y-x')}\big). \tag{7.6}$$

Proof. Since $x_i \le x_i' \le y_i$ for all $i \in \{1, \dots, n\}$, we have:

$$I_{x_i}^{y_i} = [x_i, y_i] = [x_i, x_i'] \cup [x_i', y_i] = I_{x_i}^{x_i'} \cup I_{x_i'}^{y_i} = \bigcup_{\alpha_i=0}^{1} I_{x_i+\alpha_i(x_i'-x_i)}^{x_i'+\alpha_i(y_i-x_i')},$$

and so (cf. equation (2.5) in [17, Part II]),

$$I_x^y = \prod_{i=1}^n I_{x_i}^{y_i} = \prod_{i=1}^n \Big(\bigcup_{\alpha_i=0}^{1} I_{x_i+\alpha_i(x_i'-x_i)}^{x_i'+\alpha_i(y_i-x_i')} \Big) = \bigcup_{0\le\alpha\le1} I_{x+\alpha(x'-x)}^{x'+\alpha(y-x')}.$$

The mixed difference at the left-hand side of (7.6) is given by (2.1), and again by virtue of (2.1), the mixed difference at the right-hand side of (7.6) is equal to

$$\mathrm{md}_n(f, I_{x+\alpha(x'-x)}^{x'+\alpha(y-x')}) = d\Big(\sum_{\mathrm{ev}\,\beta\le1} h(\alpha,\beta), \sum_{\mathrm{od}\,\beta\le1} h(\alpha,\beta) \Big),$$

where $h(\alpha,\beta) = f(x+(\alpha\vee\beta)(x'-x)+\alpha\beta(y-x'))$ and $\alpha\vee\beta = \alpha+\beta-\alpha\beta$. Noting that if $\alpha=\beta$, then $\alpha\vee\beta=\beta$ and $\alpha\beta=\beta$, we find

$$\begin{aligned}
\sum_{0\le\alpha\le1}\sum_{\mathrm{ev}\,\beta\le1} h(\alpha,\beta) &= \sum_{\mathrm{ev}\,\beta\le1}\sum_{\substack{0\le\alpha\le1,\\ \alpha\ne\beta}} h(\alpha,\beta) + \sum_{\mathrm{ev}\,\beta\le1} h(\beta,\beta) \\
&= \sum_{\mathrm{ev}\,\beta\le1}\sum_{\substack{0\le\alpha\le1,\\ \alpha\ne\beta}} h(\alpha,\beta) + \sum_{\mathrm{ev}\,\beta\le1} f(x+\beta(y-x)) \\
&\equiv U+u.
\end{aligned}$$

Let us show that the double sum U can be represented as

$$U = \sum_{0\ne\gamma\le1}\sum_{\substack{0\le\delta\le\gamma,\\ \delta\ne\gamma}} c_{\gamma\delta} f\big(x+\gamma(x'-x)+\delta(y-x')\big)$$

with certian integer factors $c_{\gamma\delta}$ to be evaluated below. In fact, given $0\ne\gamma\le1$ and $0\le\delta\le\gamma$ with $\delta\ne\gamma$, there exist even $\beta\le1$ and $0\le\alpha\le1$, $\alpha\ne\beta$, s.t. $\alpha\vee\beta=\gamma$ and $\alpha\beta=\delta$. In order to see this, if γ is even or δ is even, we may set $\beta=\gamma$ and $\alpha=\delta$, or $\beta=\delta$ and $\alpha=\gamma$, respectively. Now, if γ and δ are odd, then since $\delta\ne\gamma$, we can find $i\in\{1,\dots,n\}$ s.t. $\delta_i=0$ and $\gamma_i=1$, and so, if we set $\beta=(\delta_1,\dots,\delta_{i-1},1,\delta_{i+1},\dots,\delta_n)$, then $\delta\le\beta\le\gamma$, $\delta\ne\beta\ne\gamma$ and $|\beta|=|\delta|+1$ is even, and it remains to put $\alpha=\gamma+\delta-\beta$.

Given γ and δ as above, let us evaluate $c_{\gamma\delta}$. Since $\delta=\alpha\beta\le\beta\le\alpha\vee\beta=\gamma$ and, given even β, the multiindex $0\le\alpha\le1$, $\alpha\ne\beta$, s.t. $\alpha\vee\beta=\gamma$ and $\alpha\beta=\delta$, is determined uniquely by $\alpha=\gamma+\delta-\beta$, we have $c_{\gamma\delta}=|\{\text{even }\beta:\delta\le\beta\le\gamma\}|$.

In a similar manner, we find

$$\sum_{0\le\alpha\le 1}\sum_{\mathrm{od}\,\beta\le 1} h(\alpha,\beta) = \sum_{\mathrm{od}\,\beta\le 1}\sum_{\substack{0\le\alpha\le 1,\\ \alpha\ne\beta}} h(\alpha,\beta) + \sum_{\mathrm{od}\,\beta\le 1} f(x+\beta(y-x)) \equiv V+v,$$

where

$$V = \sum_{0\ne\gamma\le 1}\sum_{\substack{0\le\delta\le\gamma,\\ \delta\ne\gamma}} d_{\gamma\delta} f\big(x+\gamma(x'-x)+\delta(y-x')\big)$$

with $d_{\gamma\delta} = |\{\text{odd }\beta : \delta\le\beta\le\gamma\}|$. By Lemma 4(b), $c_{\gamma\delta} = d_{\gamma\delta}$, and so, $U = V$. Applying the translation invariance of d and inequality (5.2), we obtain inequality (7.6):

$$\begin{aligned} d(u,v) &= d(U+u,V+v) = d\Bigg(\sum_{0\le\alpha\le 1}\sum_{\mathrm{ev}\,\beta\le 1} h(\alpha,\beta), \sum_{0\le\alpha\le 1}\sum_{\mathrm{od}\,\beta\le 1} h(\alpha,\beta)\Bigg) \\ &\le \sum_{0\le\alpha\le 1} d\Bigg(\sum_{\mathrm{ev}\,\beta\le 1} h(\alpha,\beta), \sum_{\mathrm{od}\,\beta\le 1} h(\alpha,\beta)\Bigg). \qquad \square \end{aligned}$$

Remark 7.1. If $x < x' < y$ in Lemma 10, then all rectangles at the right-hand side of (7.5) are non-degenerated, i.e., $x+\alpha(x'-x) < x'+\alpha(y-x')$ for all $0\le\alpha\le 1$. Moreover, the point x' gives rise to a net partition $\{x[\sigma]\}_{\sigma=0}^{\kappa}$ of I_x^y with $x[\sigma] = (x_1(\sigma_1),\dots,x_n(\sigma_n))$ and $0\le\sigma\le\kappa$ as follows: we put $\kappa = 2 = 1+1\in\mathbb{N}^n$ and, given $i\in\{1,\dots,n\}$, we set $x_i(0) = x_i$, $x_i(1) = x_i'$ and $x_i(2) = y_i$. We note that if $0\le\sigma\le 1$, then $x[\sigma] = x+\sigma(x'-x)$, and if $1\le\sigma\le 2$, then $x[\sigma] = x'+(\sigma-1)(y-x')$. It follows that

$$I_x^y = \bigcup_{0\le\alpha\le 1} I_{x+\alpha(x'-x)}^{x'+\alpha(y-x')} = \bigcup_{1\le\sigma\le 2} I_{x+(\sigma-1)(x'-x)}^{x'+(\sigma-1)(y-x')} = \bigcup_{1\le\sigma\le\kappa} I_{x[\sigma-1]}^{x[\sigma]}.$$

However, in the general case $x\le x'\le y$ we may also have $x \not< x'$ or $x'\not< y$, and so, there exists $i\in\{1,\dots,n\}$ s.t. $x_i = x_i'$ or $x_i' = y_i$. Thus, since some coordinates of x' may be equal to the corresponding coordinates of x and/or y, certain rectangles at the right-hand side of (7.5) may degenerate into lower-dimensional rectangles, and so, by Remark 2.1, the mixed difference md_n over these rectangles is equal to zero. In order to exclude these degenerated rectangles from the consideration, we establish the following lemma.

Lemma 11. *Given $x, y \in I_a^b$ with $x < y$ and $x' \in I_x^y$, we have the following partition of I_x^y, induced by x':*

$$I_x^y = \bigcup_{\lambda \le \alpha \le \mu} I_{x+\alpha(x'-x)}^{x'+\alpha(y-x')}, \tag{7.7}$$

where the multiindices $\lambda \equiv \lambda(x, x') = (\lambda_1, \dots, \lambda_n)$ and $\mu \equiv \mu(x', y) = (\mu_1, \dots, \mu_n)$ are defined for $i \in \{1, \dots, n\}$ by

$$\lambda_i \equiv \lambda_i(x, x') = \begin{cases} 1 & \textit{if} \quad x_i = x'_i, \\ 0 & \textit{if} \quad x_i < x'_i, \end{cases} \quad \textit{and} \;\; \mu_i \equiv \mu_i(x', y) = \begin{cases} 0 & \textit{if} \quad x'_i = y_i, \\ 1 & \textit{if} \quad x'_i < y_i, \end{cases}$$

and the following inequality holds:

$$\mathrm{md}_n(f, I_x^y) \le \sum_{\lambda \le \alpha \le \mu} \mathrm{md}_n\big(f, I_{x+\alpha(x'-x)}^{x'+\alpha(y-x')}\big). \tag{7.8}$$

Proof. First, we note that, since $x_i < y_i$ for all $i \in \{1, \dots, n\}$, then $\lambda \le \mu$. In particular, if $x < x' < y$, then $\lambda = 0$ and $\mu = 1$, and we get (7.5) as a consequence of (7.7); cf. Remark 7.1.

In order to prove (7.7), given $i \in \{1, \dots, n\}$, consider the following possibilities: (i) $x'_i = x_i$ and $x'_i < y_i$; (ii) $x_i < x'_i$ and $x'_i = y_i$; and (iii) $x_i < x'_i$ and $x'_i < y_i$. We have, respectively:

(i) $\lambda_i = 1$ and $\mu_i = 1$, and so, if $\lambda_i \le \alpha_i \le \mu_i$, then $\alpha_i = 1$ and

$$I_{x_i}^{y_i} = I_{x'_i}^{y_i} = \bigcup_{\alpha_i = 1} I_{x_i+\alpha_i(x'_i-x_i)}^{x'_i+\alpha_i(y_i-x'_i)};$$

(ii) $\lambda_i = 0$ and $\mu_i = 0$, and so, if $\lambda_i \le \alpha_i \le \mu_i$, then $\alpha_i = 0$ and

$$I_{x_i}^{y_i} = I_{x_i}^{x'_i} = \bigcup_{\alpha_i = 0} I_{x_i+\alpha_i(x'_i-x_i)}^{x'_i+\alpha_i(y_i-x'_i)};$$

(iii) $\lambda_i = 0$ and $\mu_i = 1$, and so, if $\lambda_i \le \alpha_i \le \mu_i$, then $\alpha_i \in \{0, 1\}$ and

$$I_{x_i}^{y_i} = I_{x_i}^{x'_i} \cup I_{x'_i}^{y_i} = \bigcup_{\alpha_i = 0}^{1} I_{x_i+\alpha_i(x'_i-x_i)}^{x'_i+\alpha_i(y_i-x'_i)}.$$

Moreover, in all the cases (i)–(iii) the left endpoint $x_i + \alpha_i(x'_i - x_i)$ is less than the right endpoint $x'_i + \alpha_i(y_i - x'_i)$, and so, all the closed intervals above are non-degenerated. It follows that

$$I_x^y = \prod_{i=1}^{n} I_{x_i}^{y_i} = \prod_{i=1}^{n} \Bigg(\bigcup_{\lambda_i \le \alpha_i \le \mu_i} I_{x_i+\alpha_i(x'_i-x_i)}^{x'_i+\alpha_i(y_i-x'_i)} \Bigg) = \bigcup_{\lambda \le \alpha \le \mu} I_{x+\alpha(x'-x)}^{x'+\alpha(y-x')}.$$

The point x' gives rise to a net partition $\{x[\sigma]\}_{\sigma=0}^{\kappa}$ of I_x^y as follows: we put $\kappa = \mu - \lambda + 1$ and, given $i \in \{1, \ldots, n\}$, we set $x_i(0) = x_i$ and $x_i(1) = y_i$ if $\kappa_i = 1$, and $x_i(0) = x_i$, $x_i(1) = x'_i$ and $x_i(2) = y_i$ if $\kappa_i = 2$. We note that if $0 \le \sigma \le \mu - \lambda$, then $x[\sigma] = x + (\sigma + \lambda)(x' - x)$, and if $1 \le \sigma \le \kappa = \mu - \lambda + 1$, then $x[\sigma] = x' + (\sigma - 1 + \lambda)(y - x')$. Also, note that $x + \lambda(x' - x) = x$ and $x' + \mu(y - x') = y$. It follows that

$$I_x^y = \bigcup_{\lambda \le \alpha \le \mu} I_{x+\alpha(x'-x)}^{x'+\alpha(y-x')} = \bigcup_{1 \le \sigma \le \mu - \lambda + 1} I_{x+(\sigma-1+\lambda)(x'-x)}^{x'+(\sigma-1+\lambda)(y-x')} = \bigcup_{1 \le \sigma \le \kappa} I_{x[\sigma-1]}^{x[\sigma]}.$$

Now, we turn to the proof of (7.8). By Lemma 10, inequality (7.6) holds. Clearly, if $\lambda = 0$ and $\mu = 1$ (i.e., $x < x' < y$), then (7.6) implies (7.8). Assume that $\lambda \ne 0$ (i.e., $x \not< x'$) and suppose that $0 \le \alpha \le 1$ is s.t. $\lambda \not\le \alpha$. Then there exists $i \in \{1, \ldots, n\}$ s.t. $\lambda_i = 1$ and $\alpha_i = 0$, and so, $x_i = x'_i$, which implies $x_i + \alpha_i(x'_i - x_i) = x_i = x'_i = x'_i + \alpha_i(y_i - x'_i)$. It follows from Remark 2.1 that $\mathrm{md}_n(f, I_{x+\alpha(x'-x)}^{x'+\alpha(y-x')}) = 0$. Similarly, if we assume that $\mu \ne 1$ (i.e., $x' \not< y$) and suppose that $0 \le \alpha \le 1$ is s.t. $\alpha \not\le \mu$, then there exists $i \in \{1, \ldots, n\}$ s.t. $\alpha_i = 1$ and $\mu_i = 0$, and so, $x'_i = y_i$. Noting that $x_i + \alpha_i(x'_i - x_i) = x'_i = y_i = x'_i + \alpha_i(y_i - x'_i)$, we find $\mathrm{md}_n(f, I_{x+\alpha(x'-x)}^{x'+\alpha(y-x')}) = 0$. In this way inequality (7.8) follows. $\square$

Proof of Lemma 9. Suppose that $x, y \in I_a^b$, $x < y$ and $x' \in I_a^b$. We set $x'' = x + \xi(x' - x)$, where ξ is defined in (7.3) (the point x'' will play the role of x' from (7.7)). We have $x \le x'' < y$; in fact, given $i \in \{1, \ldots, n\}$, we find: if $\xi_i = 1$, then $x_i < x'_i < y_i$ and $x''_i = x'_i$ implying $x_i < x''_i < y_i$, and if $\xi_i = 0$, then $x'_i \le x_i$ or $x'_i \ge y_i$, and $x''_i = x_i$ implying $x_i = x''_i < y_i$. Applying (7.7) with x' replaced by x'', we get the following partition of I_x^y induced by x'' and, hence, by x':

$$I_x^y = \bigcup_{\lambda'' \le \alpha \le \mu''} I_{x+\alpha(x''-x)}^{x''+\alpha(y-x'')}, \tag{7.9}$$

where $\lambda'' = \lambda(x, x'')$ and $\mu'' = \mu(x'', y)$ are defined in Lemma 11, i.e., given $i \in \{1, \ldots, n\}$, we have:

$$\lambda''_i = \begin{cases} 1 & \text{if} \quad x_i = x''_i, \\ 0 & \text{if} \quad x_i < x''_i, \end{cases} \qquad \text{and} \qquad \mu''_i = \begin{cases} 0 & \text{if} \quad x''_i = y_i, \\ 1 & \text{if} \quad x''_i < y_i. \end{cases}$$

We assert that $\lambda'' = 1 - \xi$ and $\mu'' = 1$. In fact, since $x'' < y$, then $\mu'' = 1$. In order to see that $\lambda'' = 1 - \xi$, let $i \in \{1, \ldots, n\}$. If $x_i < x'_i < y_i$,

then $\xi_i = 1$, and so, $x''_i = x_i + \xi_i(x'_i - x_i) = x'_i$, which implies $x_i < x''_i$ and $\lambda''_i = 0 = 1 - \xi_i$. Now if $x'_i \le x_i$ or $x'_i \ge y_i$, then $\xi_i = 0$, and so, $x''_i = x_i$, which gives $\lambda''_i = 1 = 1 - \xi_i$.

Now, let us calculate the lower and upper indices in (7.9). We have: $x+\alpha(x''-x) = x + \alpha\xi(x' - x)$ and

$$x'' + \alpha(y - x'') = x + (1 - \alpha)\xi(x' - x) + \alpha(y - x).$$

Noting that the union in (7.9) is taken over $\alpha \le 1$ s.t. $1 - \xi \le \alpha$, we get $1 - \alpha \le \xi$, and so, $(1 - \alpha)\xi = 1 - \alpha$ implying

$$x'' + \alpha(y - x'') = x + (1 - \alpha)(x' - x) + \alpha(y - x) = x' + \alpha(y - x').$$

These calculations and observations above prove equality (7.2).

Let us show that partition (7.2) is actually induced by x'. Since $x' \in I_a^b$, by Lemma 11, the point x' induces a partition of I_a^b of the form (7.7):

$$I_a^b = \bigcup_{\lambda' \le \beta \le \mu'} I_{a+\beta(x'-a)}^{x'+\beta(b-x')},$$

where the multiindices $\lambda' = \lambda(a, x')$ and $\mu' = \mu(x', b)$ are defined in Lemma 11, i.e., given $i \in \{1, \dots, n\}$, we have:

$$\lambda'_i = \begin{cases} 1 & \text{if} \quad a_i = x'_i, \\ 0 & \text{if} \quad a_i < x'_i, \end{cases} \qquad \text{and} \qquad \mu'_i = \begin{cases} 0 & \text{if} \quad x'_i = b_i, \\ 1 & \text{if} \quad x'_i < b_i. \end{cases}$$

We assert that for each α with $1 - \xi \le \alpha \le 1$ there exists a unique $\beta \equiv \beta(\alpha)$ with $\lambda' \le \beta \le \mu'$ s.t.

$$I_{x+\alpha\xi(x'-x)}^{x'+\alpha(y-x')} = I_x^y \cap I_{a+\beta(x'-a)}^{x'+\beta(b-x')}. \tag{7.10}$$

In order to prove (7.10), we define $\beta = \beta(\alpha) = (\beta_1, \dots, \beta_n)$ by

$$\beta_i \equiv \beta_i(\alpha) = \begin{cases} \alpha_i & \text{if} \quad x'_i < y_i, \\ 0 & \text{if} \quad x'_i \ge y_i, \end{cases} \qquad i \in \{1, \dots, n\},$$

and establish equality (7.10) componentwise. Given $i \in \{1, \dots, n\}$, we consider the following two cases: (a) $x'_i < y_i$, and (b) $x'_i \ge y_i$.

In case (a) we have $\beta_i = \alpha_i$. First, assume that $x_i < x'_i$, and so, $\xi_i = 1$. It follows that if $1 - \xi_i \le \alpha_i \le 1$, then $\alpha_i = 0$ or $\alpha_i = 1$. If $\alpha_i = 0$, then we find (for $\beta_i = \alpha_i = 0$)

$$I_{x_i}^{x'_i} = [x_i, x'_i] = [x_i, y_i] \cap [a_i, x'_i] = I_{x_i}^{y_i} \cap I_{a_i+\beta_i(x'_i-a_i)}^{x'_i+\beta_i(b_i-x'_i)},$$

and if $\alpha_i = 1$, then we find (for $\beta_i = \alpha_i = 1$)

$$I_{x_i'}^{y_i} = [x_i', y_i] = [x_i, y_i] \cap [x_i', b_i] = I_{x_i}^{y_i} \cap I_{a_i+\beta_i(x_i'-a_i)}^{x_i'+\beta_i(b_i-x_i')}.$$

Now, assume that $x_i' \leq x_i$, and so, $\xi_i = 0$ and $x_i' \leq x_i < y_i \leq b_i$. It follows that if $1 - \xi_i \leq \alpha_i \leq 1$, then $\beta_i = \alpha_i = 1$ implying

$$I_{x_i}^{y_i} = [x_i, y_i] = [x_i, y_i] \cap [x_i', b_i] = I_{x_i}^{y_i} \cap I_{a_i+\beta_i(x_i'-a_i)}^{x_i'+\beta_i(b_i-x_i')}.$$

In case (b) we have $\xi_i = 0$, $\beta_i = 0$ and $a_i \leq x_i < y_i \leq x_i'$, and so, if $1 - \xi_i \leq \alpha_i \leq 1$, then $\alpha_i = 1$ and

$$I_{x_i}^{y_i} = [x_i, y_i] = [x_i, y_i] \cap [a_i, x_i'] = I_{x_i}^{y_i} \cap I_{a_i+\beta_i(x_i'-a_i)}^{x_i'+\beta_i(b_i-x_i')}.$$

Let us show that $\lambda' \leq \beta \leq \mu'$. Let $i \in \{1, \ldots, n\}$. If $a_i = x_i'$, then $\lambda_i' = 1 = \mu_i'$ and, since $x_i' < y_i$, then $\beta_i = \alpha_i$. By (7.3), $\xi_i = 0$, and so, since $1 - \xi_i \leq \alpha_i \leq 1$, then $\alpha_i = 1$, which implies $\lambda_i' = \beta_i = \mu_i'$. Now, if $x_i' = b_i$, then $\lambda_i' = 0 = \mu_i'$ and, since $x_i' \geq y_i$, then $\beta_i = 0$ (and $\xi_i = 0$), and so, $\lambda_i' = \beta_i = \mu_i'$. Finally, if $a_i < x_i' < b_i$, then $\lambda_i' = 0$ and $\mu_i' = 1$, and so, since $\beta_i \in \{0, 1\}$, then $\lambda_i' \leq \beta_i \leq \mu_i'$.

The uniqueness of $\beta(\alpha)$, for each $1 - \xi \leq \alpha \leq 1$, is a consequence of the following: if $\lambda' \leq \beta \leq \mu'$ and $\beta \neq \beta(\alpha)$, then there exists $i \in \{1, \ldots, n\}$ s.t. $\beta_i = 1 - \beta_i(\alpha)$. Arguing as in (a) and (b) above, we find that the equality (7.10) cannot hold for this β.

Now, inequality (7.4) readily follows from Lemma 11, (7.9) and (7.2). $\square$

Remark 7.2. (a) If $x' \in I_x^y$ in Lemma 9, then it is easily seen that $\xi = \mu - \lambda$, and so, equality (7.2) assumes the form:

$$I_x^y = \bigcup_{1-(\mu-\lambda)\leq\alpha\leq 1} I_{x+\alpha(\mu-\lambda)(x'-x)}^{x'+\alpha(y-x')}.$$

Although this equality looks different from (7.7), the two equalities are the same: this is verified as in (i)–(iii) of the proof of Lemma 11.

(b) If $x < x' < y$, then $\xi = 1$, $\lambda = 0$ and $\mu = 1$, and so, (7.2), (7.7) and (7.5) are identical.

(c) Here we consider a certain particular case of (7.10) and establish conditions on x', under which x' does not induce a (further) partition of I_x^y. In view of (7.10), we have:

$$I_{x+\alpha\xi(x'-x)}^{x'+\alpha(y-x')} = I_x^y \quad \text{if and only if } \xi = 0 \text{ and } \alpha = 1, \tag{7.11}$$

which is also equivalent to

$$a+\beta(x'-a)\le x \quad\text{and}\quad y\le x'+\beta(b-x') \quad\text{with}\quad \beta=\beta(1). \tag{7.12}$$

Clearly, if $\xi=0$ and $\alpha=1$, then the left-hand side equality in (7.11) holds. Conversely, if the left-hand side equality in (7.11) holds for some $1-\xi\le\alpha\le 1$, then $x+\alpha\xi(x'-x)=x$ and $x'+\alpha(y-x')=y$, and so, if we suppose that $\xi_i=1$ for some $i\in\{1,\dots,n\}$, then, by (7.3), $x_i<x_i'<y_i$, and so, $\alpha_i=0$ and $x_i'=y_i$, which is a contradiction. Thus, $\xi=0$ and $\alpha=1$.

Now, if $\xi=0$ and $\alpha=1$, then, by (7.10) and (7.11),

$$I_x^y=I_x^y\cap I_{a+\beta(x'-a)}^{x'+\beta(b-x')} \quad\text{with}\quad \beta=\beta(1), \tag{7.13}$$

which implies (7.12). Conversely, (7.12) implies (7.13), and so, the left-hand side equality in (7.11) holds, i.e., $\xi=0$ and $\alpha=1$.

This observation also shows that a point $x'\in I_a^b$ induces a 'true' partition of I_x^y provided that, for all β with $\lambda'\le\beta\le\mu'$, we have:

$$a+\beta(x'-a)\not\le x \quad\text{or}\quad y\not\le x'+\beta(b-x'),$$

which is also equivalent to $\xi\ne 0$.

Proof of Lemma 8. Let $\mathcal{P}=\{x[\sigma]\}_{\sigma=0}^{\kappa}$ for some $\kappa\in\mathbb{N}^n$ and $x'\in\mathcal{P}'$. Given $1\le\sigma\le\kappa$, we set $x_\sigma=x[\sigma-1]$, $y_\sigma=x[\sigma]$ and $\xi_\sigma(x')=\xi(x_\sigma,x',y_\sigma)$, where ξ is defined in (7.3), and note that $x_\sigma<y_\sigma$. The point x' induces a partition of $I_{x_\sigma}^{y_\sigma}=I_{x[\sigma-1]}^{x[\sigma]}$ of the form (7.2) with $x=x_\sigma$ and $y=y_\sigma$, and so, by virtue of (7.1), we get the following partition of I_a^b, induced by x':

$$I_a^b=\bigcup_{1\le\sigma\le\kappa}\ \bigcup_{1-\xi_\sigma(x')\le\alpha\le 1} I_{x_\sigma+\alpha\xi_\sigma(x')(x'-x_\sigma)}^{x'+\alpha(y_\sigma-x')}. \tag{7.14}$$

We denote by $\mathcal{P}^1$ the net partition of I_a^b corresponding to (7.14). Moreover, by (7.4), for each $1\le\sigma\le\kappa$ we have the inequality:

$$\mathrm{md}_n(f,I_{x_\sigma}^{y_\sigma})\le\sum_{1-\xi_\sigma(x')\le\alpha\le 1}\mathrm{md}_n\big(f,I_{x_\sigma+\alpha\xi_\sigma(x')(x'-x_\sigma)}^{x'+\alpha(y_\sigma-x')}\big). \tag{7.15}$$

With no loss of generality we may assume that $x'\notin\mathcal{P}$: if $x'\in\mathcal{P}$, i.e., $x'=x[\sigma']$ for some $1\le\sigma'\le\kappa$, then x' does not affect the partition $\mathcal{P}$ of I_a^b in the sense that $\mathcal{P}^1=\mathcal{P}$, and so, $\mathrm{v}_n(f;\mathcal{P}^1)=\mathrm{v}_n(f;\mathcal{P})$. In order to

see this, we note that (7.3) implies $\xi_\sigma(x') = \xi(x[\sigma-1], x[\sigma'], x[\sigma]) = 0$, and so, by Remark 7.2(c), conditions (7.11) and (7.12) hold with $\beta = \beta(1) = (\beta_1, \dots, \beta_n)$ s.t.

$$\beta_i = \begin{cases} 1 & \text{if} \quad x_i(\sigma_i') < x_i(\sigma_i), \\ 0 & \text{if} \quad x_i(\sigma_i') \geq x_i(\sigma_i), \end{cases} = \begin{cases} 1 & \text{if} \quad \sigma_i' < \sigma_i, \\ 0 & \text{if} \quad \sigma_i' \geq \sigma_i. \end{cases}$$

Summing over $1 \leq \sigma \leq \kappa$ in (7.15) and taking into account (7.1) and (7.14), we obtain the inequality

$$\mathrm{v}_n(f; \mathcal{P}) \leq \mathrm{v}_n(f; \mathcal{P}^1).$$

Replacing $\mathcal{P}$ by $\mathcal{P}^1$ in the arguments above, taking $x' \in \mathcal{P}' \setminus \mathcal{P}^1$ and denoting by $\mathcal{P}^2$ the partition of I_a^b induced from $\mathcal{P}^1$ by x', we get $\mathrm{v}_n(f; \mathcal{P}^1) \leq \mathrm{v}_n(f; \mathcal{P}^2)$. Since $\mathcal{P}' \setminus \mathcal{P}$ is a finite set, we exhaust it by points x' in a finite number of steps, arrive at the partition $\mathcal{P}'$ of I_a^b and prove the desired inequality $\mathrm{v}_n(f; \mathcal{P}) \leq \mathrm{v}_n(f; \mathcal{P}')$. □

Proof of Lemma 3. 1. First, we establish (3.5) for $\alpha = 1 = 1_n$, i.e.,

$$V_n(f, I_x^y) = \sum_{1 \leq \sigma \leq \kappa} V_n(f, I_{x[\sigma-1]}^{x[\sigma]}). \tag{7.16}$$

Modulo the notation, there is no loss of generality if we assume that $x = a$ and $y = b$, so that $\{x[\sigma]\}_{\sigma=0}^{\kappa}$ is a net partition of I_a^b.

Let $\mathcal{P}$ be an arbitrary net partition of I_a^b. Denote by $\mathcal{P}'$ the net partition of I_a^b induced from $\mathcal{P}$ by points $\{x[\sigma]\}_{\sigma=0}^{\kappa}$, so that $\mathcal{P}'$ is a refinement of $\mathcal{P}$. Given $1 \leq \sigma \leq \kappa$, set $\mathcal{P}_\sigma = \mathcal{P}' \cap I_{x[\sigma-1]}^{x[\sigma]}$ and note that $\mathcal{P}_\sigma$ is a net partition of $I_{x[\sigma-1]}^{x[\sigma]}$, and $\mathcal{P}' = \bigcup_{1 \leq \sigma \leq \kappa} \mathcal{P}_\sigma$. Then by virtue of Lemma 8, we have:

$$\mathrm{v}_n(f; \mathcal{P}) \leq \mathrm{v}_n(f; \mathcal{P}') = \sum_{1 \leq \sigma \leq \kappa} \mathrm{v}_n(f; \mathcal{P}_\sigma) \leq \sum_{1 \leq \sigma \leq \kappa} V_n(f, I_{x[\sigma-1]}^{x[\sigma]}).$$

Since $\mathcal{P}$ is arbitrary, the left-hand side in (7.16) is not greater than the right-hand side.

Let us prove the reverse inequality. If $V_n(f, I_{x[\sigma-1]}^{x[\sigma]})$ is infinite for some $1 \leq \sigma \leq \kappa$, then since $I_{x[\sigma-1]}^{x[\sigma]} \subset I_a^b = I_x^y$, the value $V_n(f, I_x^y)$ is infinite as well. Thus, we suppose that the right-hand side of (7.16) is finite. Let $\varepsilon > 0$ be

arbitrary. Given $1 \le \sigma \le \kappa$, by the definition of $V_n(f, I_{x[\sigma-1]}^{x[\sigma]})$, there exists a net partition of $I_{x[\sigma-1]}^{x[\sigma]}$, denoted by $\mathcal{P}_\sigma(\varepsilon)$, s.t.

$$\mathrm{v}_n(f; \mathcal{P}_\sigma(\varepsilon)) \ge V_n(f, I_{x[\sigma-1]}^{x[\sigma]}) - (\varepsilon/c),$$

where $c = |\{\sigma : 1 \le \sigma \le \kappa\}|$. We denote by $\mathcal{P}(\varepsilon)$ the net partition of I_a^b induced from $\{x[\sigma]\}_{\sigma=0}^{\kappa}$ by points from $\bigcup_{1\le\sigma\le\kappa} \mathcal{P}_\sigma(\varepsilon)$. Given $1 \le \sigma \le \kappa$, we set $\mathcal{P}'_\sigma(\varepsilon) = \mathcal{P}(\varepsilon) \cap I_{x[\sigma-1]}^{x[\sigma]}$ and note that $\mathcal{P}'_\sigma(\varepsilon)$ is a refinement of $\mathcal{P}_\sigma(\varepsilon)$, and $\mathcal{P}(\varepsilon) = \bigcup_{1\le\sigma\le\kappa} \mathcal{P}'_\sigma(\varepsilon)$. By virtue of Lemma 8, we find

$$\begin{aligned} V_n(f, I_a^b) &\ge \mathrm{v}_n(f; \mathcal{P}(\varepsilon)) = \sum_{1\le\sigma\le\kappa} \mathrm{v}_n(f; \mathcal{P}'_\sigma(\varepsilon)) \ge \sum_{1\le\sigma\le\kappa} \mathrm{v}_n(f; \mathcal{P}_\sigma(\varepsilon)) \\ &\ge \sum_{1\le\sigma\le\kappa} V_n(f, I_{x[\sigma-1]}^{x[\sigma]}) - \varepsilon \Bigg(\sum_{1\le\sigma\le\kappa} 1 \Bigg) \Big/ c, \end{aligned}$$

where the factor by ε is, actually, equal to 1. The desired inequality follows if we take into account the arbitrariness of $\varepsilon > 0$.

2. Now, suppose that $0 \ne \alpha \le 1$ and $\alpha \ne 1$. Note that $x\lfloor\alpha, y\lfloor\alpha \in I_{a\lfloor\alpha}^{b\lfloor\alpha}$ and $x\lfloor\alpha < y\lfloor\alpha$, and that $\{x[\sigma]\lfloor\alpha\}_{\sigma\lfloor\alpha=0}^{\kappa\lfloor\alpha}$ is a net partition of $I_{a\lfloor\alpha}^{b\lfloor\alpha}$. So, replacing $1 = 1_n$ by $1\lfloor\alpha$ (so that $|1\lfloor\alpha| = |\alpha|$) and f—by f_α^z in (7.16), we get:

$$\begin{aligned} V_{|\alpha|}(f_\alpha^z, I_x^y\lfloor\alpha) &= V_{|1\lfloor\alpha|}(f_\alpha^z, I_{x\lfloor\alpha}^{y\lfloor\alpha}) \\ &= \sum_{1\lfloor\alpha\le\sigma\lfloor\alpha\le\kappa\lfloor\alpha} V_{|1\lfloor\alpha|}\big(f_\alpha^z, I_{x[\sigma-1]\lfloor\alpha}^{x[\sigma]\lfloor\alpha}\big), \end{aligned}$$

which is equal to the right-hand side of (3.5).

This completes the proof of Lemma 3. □

8. Proof of Theorem E

Proof of Theorem E. 1. First, we show that if $x, y \in I_a^b$, $x < y$, and $0 \ne \alpha \le 1$, then

$$\mathrm{md}_{|\alpha|}(f_\alpha^a, I_x^y\lfloor\alpha) = \lim_{j\to\infty} \mathrm{md}_{|\alpha|}((f_j)_\alpha^a, I_x^y\lfloor\alpha). \tag{8.1}$$

By virtue of (3.3), we have:

$$\mathrm{md}_{|\alpha|}(f_\alpha^a, I_x^y\lfloor\alpha) = d\Bigg(\sum_{\mathrm{ev}\,\theta\le\alpha} f(\underbrace{a + \alpha(x-a) + \theta(y-x)}_{(\cdots)}), \sum_{\mathrm{od}\,\theta\le\alpha} f(\cdots) \Bigg),$$

and a similar equality holds for f_j in place of f. Applying the inequalities $|d(u,v) - d(u',v')| \le d(u,u') + d(v,v')$, $u,v,u',v' \in M$, and (5.2) and taking into account the pointwise convergence of f_j to f, we find

$$\begin{aligned}
&\Big|\mathrm{md}_{|\alpha|}((f_j)^a_\alpha, I^y_x\lfloor\alpha) - \mathrm{md}_{|\alpha|}(f^a_\alpha, I^y_x\lfloor\alpha)\Big| \\
&\quad\le d\Bigg(\sum_{\mathrm{ev}\,\theta\le\alpha} f_j(\cdots), \sum_{\mathrm{ev}\,\theta\le\alpha} f(\cdots)\Bigg) + d\Bigg(\sum_{\mathrm{od}\,\theta\le\alpha} f_j(\cdots), \sum_{\mathrm{od}\,\theta\le\alpha} f(\cdots)\Bigg) \\
&\quad\le \sum_{\mathrm{ev}\,\theta\le\alpha} d(f_j(\cdots), f(\cdots)) + \sum_{\mathrm{od}\,\theta\le\alpha} d(f_j(\cdots), f(\cdots)) \\
&\quad= \sum_{0\le\theta\le\alpha} d(f_j(\cdots), f(\cdots)) \to 0 \quad \text{as} \quad j\to\infty.
\end{aligned}$$

2. In the rest of this proof we need only the inequality

$$\mathrm{md}_{|\alpha|}(f^a_\alpha, I^y_x\lfloor\alpha) \le \liminf_{j\to\infty} \mathrm{md}_{|\alpha|}((f_j)^a_\alpha, I^y_x\lfloor\alpha), \tag{8.2}$$

which readily follows from (8.1) and is applied one more time in the proof of Theorem 2 (Step 5). If $\mathcal{P} = \{x[\sigma]\}_{\sigma=0}^{\kappa}$ is a net partition of I^b_a, then $\mathcal{P}\lfloor\alpha = \{x[\sigma]\lfloor\alpha\}_{\sigma\lfloor\alpha}^{\kappa\lfloor\alpha}$ is net partition of $I^b_a\lfloor\alpha$, and so, given $1\le\sigma\le\kappa$, setting $x = x[\sigma-1]$ and $y = x[\sigma]$ in (8.2), we find

$$\begin{aligned}
\sum_{1\lfloor\alpha\le\sigma\lfloor\alpha\le\kappa\lfloor\alpha} \mathrm{md}_{|\alpha|}(f^a_\alpha, I^{x[\sigma]}_{x[\sigma-1]}\lfloor\alpha) &\le \sum_{1\lfloor\alpha\le\sigma\lfloor\alpha\le\kappa\lfloor\alpha} \liminf_{j\to\infty} \mathrm{md}_{|\alpha|}((f_j)^a_\alpha, I^{x[\sigma]}_{x[\sigma-1]}\lfloor\alpha) \\
&\le \liminf_{j\to\infty} \sum_{1\lfloor\alpha\le\sigma\lfloor\alpha\le\kappa\lfloor\alpha} \mathrm{md}_{|\alpha|}((f_j)^a_\alpha, I^{x[\sigma]}_{x[\sigma-1]}\lfloor\alpha) \\
&\le \liminf_{j\to\infty} V_{|\alpha|}((f_j)^a_\alpha, I^{x[\sigma]}_{x[\sigma-1]}\lfloor\alpha).
\end{aligned}$$

By the arbitrariness of $\mathcal{P}$, we infer that

$$V_{|\alpha|}(f^a_\alpha, I^{x[\sigma]}_{x[\sigma-1]}\lfloor\alpha) \le \liminf_{j\to\infty} V_{|\alpha|}((f_j)^a_\alpha, I^{x[\sigma]}_{x[\sigma-1]}\lfloor\alpha).$$

We conclude that

$$\begin{aligned}
\mathrm{TV}(f, I^b_a) &= \sum_{0\ne\alpha\le 1} V_{|\alpha|}(f^a_\alpha, I^b_a\lfloor\alpha) \le \sum_{0\ne\alpha\le 1} \liminf_{j\to\infty} V_{|\alpha|}((f_j)^a_\alpha, I^b_a\lfloor\alpha) \\
&\le \liminf_{j\to\infty} \sum_{0\ne\alpha\le 1} V_{|\alpha|}((f_j)^a_\alpha, I^b_a\lfloor\alpha) = \liminf_{j\to\infty} \mathrm{TV}(f_j, I^b_a). \qquad \square
\end{aligned}$$

References

[1] C. R. Adams, J. A. Clarkson, Properties of functions $f(x,y)$ of bounded variation, Trans. Amer. Math. Soc. 36 (4) (1934) 711–730.

[2] L. Ambrosio, Metric space valued functions of bounded variation, Ann. Scuola Norm. Sup. Pisa Cl. Sci. 17 (1990) 439–478.

[3] L. Ambrosio, N. Fusco, D. Pallara, Functions of Bounded Variation and Free Discontinuity Problems, Clarendon Press, Oxford, 2000.

[4] P. Antosik, The investigation of continuity of a function of several variables, Ann. Soc. Math. Polon., Ser. I: Comment. Math. 10 (1966) 101–104.

[5] M. Balcerzak, S. A. Belov, V. V. Chistyakov, On Helly's principle for metric semigroup valued BV mappings of two real variables, Bull. Austral. Math. Soc. 66 (2) (2002) 245–257.

[6] V. Barbu, Th. Precupanu, Convexity and Optimization in Banach Spaces, second ed., Reidel, Dordrecht, 1986.

[7] S. A. Belov, V. V. Chistyakov, A selection principle for mappings of bounded variation, J. Math. Anal. Appl. 249 (2) (2000) 351–366.

[8] P. Billingsley, Convergence of Probability Measures, John Wiley and Sons, New York, London, 1968.

[9] S. Bochner, Lectures on Fourier Integrals (with an author's supplement on monotonic functions, Stieltjes integrals, and harmonic analysis), Annals of Math. Studies, No. 42, Princeton University Press, Princeton, N.J., 1959.

[10] V. V. Chistyakov, On the theory of multivalued mappings of bounded variation of one real variable, Sbornik Math. 189 (5–6) (1998) 797–819.

[11] V. V. Chistyakov, Superposition operators in the algebra of functions of two variables with finite total variation, Monatsh. Math. 137 (2) (2002) 99–114.

[12] V. V. Chistyakov, Functions of several variables of finite variation and superposition operators, in: Real Analysis Exchange 26th Summer Symposium, Lexington, VA, USA, 2002, pp. 61–66.

[13] V. V. Chistyakov, Metric semigroups and cones of mappings of finite variation of several variables, and multivalued superposition operators, Dokl. Math. 68 (3) (2003) 445–448.

[14] V. V. Chistyakov, A selection principle for mappings of bounded variation of several variables, in: Real Analysis Exchange 27th Summer Symposium, Opava, Czech Republic, 2003, pp. 217–222.

[15] V. V. Chistyakov, Selections of bounded variation, J. Appl. Anal. 10 (1) (2004) 1–82.

[16] V. V. Chistyakov, Abstract superposition operators on mappings of bounded variation of two real variables. I and II, Siberian Math. J. 46 (3) (2005) 555–571, and 46 (4) (2005) 751–764.

[17] V. V. Chistyakov, A Banach algebra of functions of several variables of finite total variation and Lipschitzian superposition operators. I and II, Nonlinear Anal. 62 (3) (2005) 559–578, and 63 (1) (2005) 1–22.

[18] V. V. Chistyakov, The optimal form of selection principles for functions of a real variable, J. Math. Anal. Appl. 310 (2) (2005) 609–625.

[19] V. V. Chistyakov, A pointwise selection principle for functions of one variable with values in a uniform space, Siberian Adv. Math. 16 (3) (2006) 15–41.

[20] V. V. Chistyakov, C. Maniscalco, A pointwise selection principle for metric semigroup valued functions, J. Math. Anal. Appl. 341 (1) (2008) 613–625.

[21] V. V. Chistyakov, C. Maniscalco, Yu. V. Tretyachenko, Variants of a selection principle for sequences of regulated and non-regulated functions, in: Topics in Classical Analysis and Applications in Honor of Daniel Waterman (L. De Carli, K. Kazarian and M. Milman, Eds.) World Scientific Publishing, Singapore, 2008, pp. 45–72.

[22] J. A. Clarkson, C. R. Adams, On definitions of bounded variation for functions of two variables, Trans. Amer. Math. Soc. 35 (1933) 824–854.

[23] F. S. De Blasi, On the differentiability of multifunctions, Pacific J. Math. 66 (1) (1976) 67–81.

[24] E. Giusti, Minimal Surfaces and Functions of Bounded Variation, Birkhäuser, Boston, 1984.

[25] C. Goffman, T. Nishiura, D. Waterman, Homeomorphisms in Analysis, Math. Surveys Monogr., vol. 54, Amer. Math. Soc., Providence, RI, 1997.

[26] G. H. Hardy, On double Fourier series, and especially those which represent the double zeta-function with real and incommensurable param-

eters, Quart. J. Math. Oxford 37 (1905/06) 53–79.

[27] E. Helly, Über lineare Funktionaloperationen, Sitzungsber. Naturwiss. Kl. Kaiserlichen Akad. Wiss. Wien 121 (1912) 265–297.

[28] T. H. Hildebrandt, Introduction to the Theory of Integration, Academic Press, New York, London, 1963.

[29] D. Idczak, S. Walczak, On Helly's theorem for functions of several variables and its applications to variational problems, Optimization 30 (1994) 331–343.

[30] L. D. Ivanov, Variations of Sets and Functions, Nauka, Moscow, 1975.

[31] B. Lenze, On the points of regularity of multivariate functions of bounded variation, Real Anal. Exchange 29 (2) (2003/04) 646–656.

[32] A. S. Leonov, On the total variation for functions of several variables and a multidimensional analog of Helly's selection principle, Math. Notes 63 (1) (1998) 61–71.

[33] J. J. Moreau, Bounded variation in time, in: Topics in Nonsmooth Mechanics (J. J. Moreau, P. D. Panagiotopoulos, G. Strang, eds.), Birkhäuser-Verlag, Basel, 1988, pp. 1–74.

[34] N. S. Nadirashvili, Helly's principle of choice for functions of two variables, Moscow University Math. Bull., Ser. Math. Mech., 3 (1975) 3–10.

[35] M. Picone, Sulle variazione totale di una funzione metrica, Rend. Sem. Mat. Fis. Milano 30 (1960) 59–92.

[36] H. Rådström, An embedding theorem for spaces of convex sets, Proc. Amer. Math. Soc. 3 (1952) 165–169.

[37] L. Schwartz, Analyse Mathématique, vol. 1, Hermann, Paris, 1967.

[38] G. Vitali, Sulle funzione integrali, Atti Accad. Sci. Torino Cl. Sci. Fis. Mat. Natur. 40 (1904/05) 1021–1034; and Opere sull'analisi reale, Cremonese (1984) 205–220.

[39] A. G. Vituškin, On Multidimensional Variations, Gostekhizdat, Moscow, 1955.

[40] W. H. Young, G. C. Young, On the discontinuities of monotone functions of several variables, Proc. London Math. Soc. 22 (1924) 124–142.

[41] W. P. Ziemer, Weakly Differentiable Functions. Sobolev Spaces and Functions of Bounded Variation, Graduate Texts in Math., vol. 120, Springer, New York, 1989.